Probabilistic Modelling of the Evolution of Ecological Interaction Networks

by

Henintsoa Onivola Minoarivelo

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science at Stellenbosch University

Department of Mathematical Sciences,
Computer Science Division,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Prof Konrad Sheffler
Co-Supervisor: Dr Cang Hui

September 2011

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: .....
H. O. Minoarivelo

Date: September 15, 2011 .....

Abstract

Probabilistic Modelling of the Evolution of Ecological Interaction Networks

H. O. Minoarivelo

Department of Mathematical Sciences,
Computer Science Division,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MSc

September 2011

In any ecological system, organisms need to interact with each other for their survival. Such interactions form ecological networks which are usually very complex. Nevertheless, they exhibit well defined patterns; these regularities are often interpreted as products of meaningful ecological processes. As the networks are evolving through time, biological evolution is one of the factors that affects ecological network architecture. In this work, we develop a mathematical model that represents the evolution through time of such ecological interaction networks. The problem is approached by modelling network evolution as a continuous time Markov process, in such a way that the interactions in which a parent species is involved are potentially inherited by its descendant species. This approach allows us to infer ecological parameters and ecological network histories from real-world network data, as well as to simulate ecological networks under our model. While ecologists have long been aware of the influence of evolutionary processes in shaping ecological networks, we are now able to evaluate the importance of such influence.

Uittreksel

Probabilistic Modelling of the Evolution of Ecological Interaction Networks

("Probabilistiese Modellering van die Ewolusie van Ekologiese Interaksienetwerke ")

H. O. Minoarivelo

Departement Wiskundige Wetenskappe,
Afdeling Rekenaarwetenskap,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MSc

September 2011

In enige ekologiese stelsel benodig organismes wisselwerkings met mekaar ten einde te oorleef. Sulke interaksies vorm ekologiese netwerke wat gewoonlik baie kompleks is maar nogtans goed-gedefinieerde patrone vertoon. Hierdie patrone word dikwels geïnterpreteer as die produk van betekenisvolle ekologiese prosesse. Aangesien die netwerke met die verloop van tyd ontwikkel, is biologiese ewolusie een van die faktore wat ekologiese netwerkargitektuur beïnvloed. In hierdie studie ontwikkel ons 'n wiskundige model wat die ewolusie van sulke ekologiese interaksienetwerke voorstel. Die probleem word benader deur netwerkewolusie as 'n kontinue-tyd Markov-proses te modelleer, op so 'n manier dat die interaksies waarin 'n voorouerspesie betrokke is potensieel oorerf kan word deur die afstammelingspesies. Hierdie benadering laat ons toe om ekologiese parameters en ekologiese netwerkgeskiedenisse vanuit regte-wêreld data af te lei, sowel as om ekologiese netwerke onder ons model te simuleer. Alhoewel ekoloë al lank reeds bewus is van die invloed wat ewolusionêre prosesse het op

die vorming van ekologiese netwerke, is ons nou in staat om die belangrikheid van hierdie invloed te evalueer.

Acknowledgements

I thank God, My Heavenly Father, for His guidance and love, making this thesis possible.
My deepest acknowledgment goes to Prof Konrad Sheffler for his diligent and patient guidance, his invaluable feedback throughout this thesis, and his help in so many aspects of this thesis. I am also very grateful to him for having translated the abstract into Afrikaans.
I owe a special debt to Dr Cang Hui and Dr John Terblanche for those enjoyable and inspiring sessions of discussion and idea exploration, and to Prof Sergei Kosakovsky Pond for his help with the model implementation.
My thanks also go to all my friends and colleagues for their support and encouragement. A heartily appreciation goes to my family for their prayers and loyal support during the good and the hard times. I am exceptionally grateful to Ando for his devoted support and help throughout the accomplishment of this thesis.
Finally, the financial assistance of the Faculty of Science of the University of Stellenbosch, the African Institute for Mathematical Sciences (AIMS) and the Center of Excellence for Invasion Biology (CIB) is hereby acknowledged.

Dedications

To my parents

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to the Problem

As a rule, natural systems are very complex. Understanding their functioning may sometimes require more than empirical observations. The use of theoretical and analytical tools in solving problems in ecology has existed for a long time, but only recently has there been an explosion in the field due to the acquirement of very large datasets and the recent increase in computational power allowing large scale calculations and simulations. Theory is not only about a representation of the real but it also gives a framework for analyses and predictions. The study of interactions between organisms holds an important place in ecology because the survival of species depends largely on their interactions with their environment. Networks formed by those interactions evolve through time: while new species become involved in the network, other species may no longer be a network participant. The absence or the presence of a species in the interaction may primarily depend on the evolution of that species.

In this thesis, we propose to mathematically model the evolution of such ecological networks through time. For this purpose, we make use of a probabilistic approach in which species gain or lose interactions at a fixed rate through evolutionary time. We assume that when a parent species gives rise to a pair of descendant species, the descendant species inherit the interactions of the parent. This makes sense because shared morphological and physiological traits between parents and their descendants ensure a high propensity for them to interact with the same partner.

Most ecological network studies focus on exploring the structure of the networks. This structure depends on many factors. One of them is the past evolutionary history of the

species involved in the networks. Although the fact that the structure of ecological networks is influenced by evolutionary history has already been acknowledged in the field of ecology, few works have been done in quantifying this influence. By building an ecological network model that incorporates the process of evolution itself, we aim to quantify the extent to which ecological network structure is governed by evolution, and this will be done by means of simulations under our model. Our model is an attempt to use mathematical, computational and phylogenetic tools towards solving a problem in evolutionary ecology.

1.2 Thesis Layout

In addition to this introductory chapter, this thesis comprises 5 further chapters. Chapter 2 puts the reader in the context of the thesis. It gives a global view of complex networks with a focus on ecological networks. We will also list, to the best of our knowledge, works already done on network models which are related to our model.
Chapter 3 aims to give information on how to perform a standard phylogenetic inference, because some of the phylogenetic tools described in this chapter will be used in building our model.
The model itself will be described in chapter 4. We will explain how we constructed our model by detailing the method we used. We will also describe how we performed simulations under the model. Those simulations were performed in order to look at the effect of the phylogenetic history of the species on ecological network structure.
In chapter 5, we will state the results we obtained from simulations. By interpreting those results, we will also try to give answers to fundamental questions raised in this thesis.
Finally, the thesis closes with a summary of the contributions we made as well as recommendations one has to be aware of when using our model. Indication of possible future avenues which deserve to be investigated will also be given.

Chapter 2

Literature Review

In a society and in nature, there is a large range of systems that do not present a trivial structure. They are often modelled by complex networks and form a field of study of their own. Complex network studies have always been associated with the field of graph theory which allows an abstract representation of networks by nodes linked (or not) by edges. The word 'network' normally invokes interactions between the entities forming the network. Da Fontoura Costa et al. [2011] present a survey of the applications of complex networks in a diversity of domains such as biological networks, social networks, ecological networks and the Internet. They gave an overview of the observed structure of networks as well as a description of models used in each of these domains.
In the centre of interaction network studies is the analysis of their characteristics and their dynamics. Interaction network characteristics have been the main focus in the field, but the latter, describing the evolution of the networks over time, also deserves attention since it can give explanations for the former. The first section of this chapter will discuss the structure of complex networks. It will be followed by a section on models of evolving networks where both structural models and models of evolution will be discussed. Since ecological networks are our central concern, the final section will be devoted to them.

2.1 On the Structure of Complex Networks

In order to characterize complex networks, they first need to be measured. We start by listing the most commonly used measurements of complex networks in the field of graph theory.

2.1.1 Measurements

Node degree

The degree of a node is the total number of connections shared by the node. The average node degree in a network gives the connectivity of the network. Another measurement related to the connectivity is the node degree distribution that gives the probability $P(k)$ of a uniformly chosen node to have a degree k.

Clustering Coefficient

The clustering coefficient is a measure of how the nodes in the network tend to cluster together. Given that the neighbours of a node are the nodes directly connected to that node, the clustering coefficient of a node is the ratio between the number of connections among the neighbours of the node and the maximum possible number of connections among these neighbours. Usually, to characterize the entire network, the average over all the nodes of the clustering coefficients is computed.

Path Length

The length of a path connecting two nodes is the total number of edges along the path. The shortest path is usually considered as a measurement of the distance between two nodes. The whole network can be characterized by the average path length, which is the average value of the shortest path lengths between all pairs of nodes in the network.

Strength

Sometimes, instead of just considering the presence or the absence of a link between two nodes, a weight is assigned to each edge, forming a weighted network. The strength of a node is the sum of the weights of the edges that connect the node.

Assortativity

Assortativity refers to the fact that similar nodes are connected. The similarity of the nodes can be evaluated in terms of any measurement on the nodes. Commonly used approaches are assortativity in degree and assortativity in strength. Assortativity in degree is measured

by examining the correlation between the degree of each node and the average degree of the the neighbours of the node. A positive correlation indicates an assortativity while a negative correlation indicates a disassortativity.

2.1.2 Characteristics

Several real-world networks, such as some technological, social or biological networks, possess the small-world feature [Watts and Strogatz, 1998]. A small-world network is characterized by the fact that most nodes are connected by a short path. These networks have intermediate properties between a randomly structured network and a network with a regular structure. They exhibit a high clustering coefficient which denotes a dense connectivity between neighbours of a given node forming a module and sparser external connections to other modules. They also exhibit a small average path length like random graphs.
The exploration of several large databases describing the structure of real-world networks spanning diversified fields such as the World Wide Web (WWW) or the citation patterns in science allowed Barabási and Albert [1999] to report a scale-free power-law in the node degree distribution. That is to say, $P(k) \propto k^{-\gamma}$ for large values of k. The exponent γ often lies in the range $2 \leq \gamma \leq 4$.
In addition to the scale-free feature observed in the degree distribution, both the connection weights and the strength of nodes in many real-world complex networks also obey a power-law distribution [Li and Chen, 2004].
In particular, many real biological networks, such as gene regulatory, protein-protein interaction and metabolic networks are known to have a hierarchical structure of modules. This property is highlighted by the fact that if we plot the clustering coefficient of each node against the degree of the node, a power-law behaviour is observed. Several biological networks also exhibit a disassortativity in degree [Takemoto and Oosawa, 2006].

2.2 Evolving Network Models: an Overview

Almost all real-world networks evolve through time: some nodes and connections may disappear while new nodes and links may appear. While some evolving network models emphasize the modelling of the dynamics of networks, some models are built in order to find plausible explanations of the empirically observed structures of complex networks. We first discuss structural models, followed by models that are similar to ours in that they explicitly incorporate evolution.

2.2.1 Models Based on Structure

The main purpose of evolving network models based on network structures is to have rules governing the network dynamics that will result in networks that have the empirically observed structure.

Scale-free Network Models

The first network model able to reproduce the scale-free property of node degree distribution was proposed by Barabási and Albert [1999]. They explained that this scale-free behaviour is a consequence of two features of real-life networks: the growth rule and the preferential attachment rule. In fact, real networks are observed to grow through time by addition of new nodes. These new nodes are connected to already existing nodes according to a preferential connectivity rule: they have higher probability to connect to a node having a large number of connections than to poorly connected nodes. Based on those two assumptions, they constructed a model of network evolution reproducing a scale-free behaviour of the node degree distribution at any time step.
The construction rules in the Barabási-Albert model (B-A model) was then later considered as basis in other models.
In this trend, Krapivsky et al. [2000] presented a growing network model that generalizes the B-A model. They proposed the form of the probability of a new node to connect to a pre-existing node with k pre-existing links, A_k to be proportional to k^ν ($\nu \geq 0$). By assuming this form of A_k, they retained the preferential attachment rule of the B-A model. They then examined the possible structure of the network for different values of the exponent ν. For $\nu < 1$, the node degree distribution N_k is exponential. $\nu > 1$ shows an extreme structure where a single node connects to nearly every node in the graph. Only the case when $\nu = 1$ leads to a power-law distribution: $N_k \propto k^{-\gamma}$. Their generalized approach suits not only the special and dominant case of a power-law distribution in degree, but all networks that grow through time.
Li and Chen [2003] tried to overcome the limitation of the B-A model that it has a fixed exponent in the power-law degree distribution while the empirically measured exponents vary. They built a model producing a node degree distribution that represents a transition between a power-law scaling and an exponential distribution, by using the concept of 'local-world network': a new added node may only connect to a randomly selected set of nodes (its 'local world') following the preferential attachment rule.
Instead of just assuming the existence or absence of the connections, Li and Chen [2004]

built a model that can represent weighted links between nodes. They used the preferential attachment rule and a preferential strengthening rule as well. These processes allowed them to have a model that mimics the three observed power-law behaviours: in the degree distribution, in the strength of nodes and in the connection weights.
Li and Maini [2005] elaborated an evolving network model with a community structure considering the two rules used in the B-A model. On the one hand, they retained the growth rule by adding a new node at each time step. On the other hand, they defined two rules for the preferential attachment: an inner-community preferential attachment rule describing the probability of the nodes in the same community to connect with a new node, and an inter-community preferential attachment rule stating the probability of nodes in other communities to connect with the new node. The communities are defined as sets of densely connected nodes, as in the case of modules. Their model also reproduced a power-law distribution of the node degree.

Small-world Models

The small-world model was proposed by Watts and Strogatz [1998]. In their model, an edges rewiring process is used: they began by arranging the nodes in a ring structure and connecting each node to its k nearest neighbours.This results in a graph with a regular form. Next, each of the edges in this graph is removed and reconnected with probability p to a randomly chosen node in the network. The edges connecting a node to its nearest neighbours are considered first, then the edges connecting each node to its second nearest neighbour, etc, until all the edges in the initial regular graph have been rewired. The resulting network has properties between a regular graph and a random graph, and displays the small-world phenomenon.
Other models were inspired by both the small-world phenomenon and the B-A model.
That is the case for a model built by Jost and Joy [2002]. A distance preference function characterizes their model: the probability of the formation of a new link depends on the distance separating the two linked nodes. For instance, one can assume that the connection between two separated nodes by a distance of 2 (neighbour of neighbour) is the most probable link to be formed. In the network construction, when a new node is added, it is first connected randomly to an existing node. The formation of links connecting it to the other nodes depends on the distance preference function. They found that when the distance preference function is such that the shortest distance is always preferred, one ends up with a network with high clustering coefficient, short average path length and with the scale-free behaviour of the degree distribution. This means that their model possesses at the same

time the properties of the small-world model and the B-A model.
In biology, Takemoto and Oosawa [2006] also proposed such a network model. They modified the growth rule of the B-A model: instead of adding a single node at each time step, an entire module of constant size is added. Then, to merge the new module with the pre-existing nodes, they followed a preferential attachment rule combined with a fitness-driven mechanism: the probability of a node i to get connected to a new node is proportional to its degree k_i (preferential attachment rule) and its fitness f_i (fitness-driven mechanism). Initialized to 0 for all nodes, the fitness value f_i is increased at each step if the node i acquires a new link. The resulting model mimics the scale-free distribution of node degree, the small-world property and the hierarchical structure of modules. Additionally, the inclusion of the fitness parameter contributes to a disassortativity in degree.

2.2.2 Models of evolution

For some evolutionary network models, the emphasis is not in reproducing the network structure observed empirically but in modelling the process of evolution of the network.
In this context, Borgnat et al. [2008] indicate recently used approaches to characterize the dynamics of complex networks. According to them, evolving networks can be considered as a sequence of snapshots at different times. Each snapshot can be studied as a static network. The evolution of some global parameters such as the number of nodes or the number of links can then be analysed by considering their states at each snapshot. They also stated a more local approach: dynamic properties such as birth, death, or growth can be assigned to a specific group of nodes or even to each individual node.
Grindrod and Higham [2010] used the snapshot approach to develop an evolving network model based on a discrete time Markov chain. Their model is adapted for networks in which a spatial location can be assigned to each node. For example, in a social network, the physical location of each individual can be considered as its spatial location in the graph representation. For such networks, a distance can then be assigned between two nodes: this distance is called the 'range' separating the two nodes. In their model, the formation of a link between two nodes depends probabilistically on that 'range'. Each edge is assigned a weight which represents the 'range' separating the two connected nodes. The range parameter is inferred from the snapshots of real data for all possible edges. They showed that this range-dependent model is applicable for real-life networks as in telecommunications. Subsequently, they generalized their model, but this time considering other extra parameters on which the probability of the birth or the death of a link depends, such as local clustering or the degree of nodes [Grindrod and Higham, 2011].

In the modelling of biological networks, a completely different approach was proposed by Vázquez et al. [2003]. They modelled a protein-protein interaction network based on the evolutionary history of the proteins: the duplication-divergence model. The model works as follows: during a duplication step, a node, considered as a duplicate of a randomly selected node, is created. Then, this new node shares all the connections of its duplicate, and connects with it according to a certain probability. Then comes the divergence step: among all the proteins to which the two previous proteins are connected, one node is removed with a certain probability. The two processes are repeated at each time step. They showed that this model captures some features of a real protein interaction network, such as the power-law behaviour of the node degree, or the number of proteins linked into triplets or quartets.
Although approaches and methods are very diversified in the modelling of an evolving network, they are all based on a probabilistic approach: the addition of a node or the creation of a new connection are often defined to follow a certain probability. Such a probabilistic approach is also considered in the model we develop in this thesis. However, most of the above models focus on reproducing the characteristics of empirical networks. Consequently, the rules governing the evolution are based on suppositions only. By doing so, modelling the real processes of evolution can be neglected. For instance, most of the models only consider the growth of the network by addition of new nodes and neglect the fact that at some time, some nodes may disappear. This issue is overcome in our model since it is based on the phylogenetic history of the set of interacting individuals. The model of protein-protein interaction [Vázquez et al., 2003] is similar to our model in the sense that in both cases the history of the evolution is modelled with reference to divergence and duplication or speciation. However, our model is simpler.

2.3 Ecological Interaction Networks

Every species interacts with its environment and other species, forming a complex ecological network. This interaction is fundamental for the survival of the species. Thus ecological interaction networks ensure the functioning of the entire ecosystem. That is why ecological interaction studies have always played a central role in ecology.

2.3.1 Classification

Ecological interactions can be classified according to the way species interact: we can distinguish mutualistic, commensalistic, amensalistic and antagonistic interactions.

2.3.1.1 Mutualistic Interaction Networks

A mutualistic interaction occurs when both of the two interacting organisms benefit from the association. A common example of a mutualistic network is a plant-pollinator network in which the pollinator (such as bees or birds) drinks the nectar of the plant species. In return, after being covered by pollen from that plant, the pollinator transfers the pollen to the corresponding female plant, ensuring the pollination process in that way. A frugivory network is an example of a mutualistic network as well; the frugivore animal eats the fruit and in return disperses the seed of the fruit. This allows the plant to have progeny. Compared to other mutualistic interactions, interactions between plants and their pollinators or frugivores have gained a particular attention for ecologists (for example, see: Levey et al. [2002], Landry [2010]). Another example of a mutualistic interaction is the relationship between humans and their intestinal bacteria. The bacteria benefit by getting food that the human cannot digest while the human benefits by being able to finish the digestion. Boucher et al. [1982] have identified two types of mutualistic interactions: direct mutualism in which the two species interact physically and indirect mutualism in which the two species benefit from each other without necessary being in a direct contact. Interesting reviews on mutualistic interactions have been made by Boucher et al. [1982] and Bronstein [1994].

2.3.1.2 Commensalistic Interaction Networks

A commensalism is an interaction between two organisms in which one organism benefits and the other neither benefits nor is harmed. In many cases, the beneficiary species uses the host species as a home and/or a transportation. For example, in order to get more air and light, some tropical orchids grow on the support of a tree without harming or benefiting the tree. Another example is the case of barnacles, a sedentary crustacean species. To gain support, they must attach themselves to a solid substance such as the shell of a scallop, leaving the scallop unaffected.

2.3.1.3 Amensalistic Interaction Networks

One can talk about an amensalistic interaction when one organism is affected by the interaction while the other one derives no benefit. Usually, this occurs when one organism is harmed by a product of the other organism. That is for instance the case when sheep make trails in grass that they trample on. In this manner, the grass can be killed. Another example is the

black walnut tree that secretes a chemical substance which can kill some species of plants in their neighbourhood.

2.3.1.4 Antagonistic Interaction Networks

When the relationship between two organisms reflects an opposition or a competition resulting in one organism benefiting at the expense of the other one, the interaction is called antagonism. One can consider the example of a predator-prey interaction network (food web) formed by a group of species that feed on each other. Food-web studies have long been central to ecological research, consequently attract many researchers (for example, see: Pimm [1980], Williams and Martinez [2000]). Another example of an antagonistic network is a parasitism interaction network in which parasites, living on or inside the host, benefits while hosts are harmed.

2.3.1.5 Competition

When two organisms share a common requirement for a limited resource, they are engaged in a competition. Competition is a mutually detrimental interaction between the two organisms. A competition may occur between organisms of the same species (intraspecific competition) or between organisms of two separate species (interspecific competition). An example of competition is when two trees grow close together: they will compete for nutrients in the soil.

Regarding the effect of the interaction on the pair of organisms, the way ecological interaction networks are classified can be summarized in table 2.1.

Table 2.1: Classification of ecological interactions

species A: / species B:	wins	is not affected	lose
wins	**Mutualism**	**Commensalism**	**Antagonism**
is not affected	**Commensalism**	$(\star)$	**Amensalism**
lose	**Antagonism**	**Amensalism**	**Competition**

$(\star)$ represents the uninteresting case where the interaction is neutral for both of the species.

In reality, the classification of an ecological interaction into these categories remains ambiguous since it also depends on how far the interaction has been investigated in term of details and time. For instance, a commensalistic interaction can become an antagonism if the host is shown to be negatively affected by the relationship in the long term.
Historically, ecological network studies have tended to focus mainly on antagonistic networks, specifically on food webs. Other types of interactions, such as mutualism, have only been explored recently. Our study will be concentrated on mutualistic networks.

2.3.2 Structure of Mutualistic Networks

Assumptions on the structure of ecological networks result from immeasurable hours of careful observation of nature by means of field work followed by statistical analysis. Like other types of real-world network, ecological networks have been found to have a non-random structure. Therefore, the study of the architecture of mutualistic networks has become a central focus. For reviews on the subject, see Bascompte and Jordano [2007], Vázquez et al. [2009].

2.3.2.1 Degree Distribution

Early studies on the relative abundance of specialists (species interacting with only a few species) and generalists (species that can make use of different varieties of resources, thus interacting with many other species) motivated the examination of the degree distribution in an ecological network. In fact, Waser et al. [1996] examined the level of generalization of pollination webs and concluded that, in general, a pollinator uses several plant species and a plant species is visited by several pollinators. This makes generalization a rule rather than an exception. Some time later, Memmott [1999] explored methods usually applied for food webs to study the structure of plant-pollinator webs and supported Waser et al. [1996] on the domination of generalist species. Vázquez and Aizen [2003] contradicted the assumption of Waser et al. [1996] by making predictions on the patterns of generalization in plant-pollinator interactions on the basis of a null model. Thus, they observed a high number of both the generalists and the specialists compared to the null model expectation.
Those earlier studies only gave a general approach of how abundantly species interact in mutualistic networks, without suggesting an appropriate distribution for the number of interactions per species. Consequently, Jordano et al. [2003] proposed to fit the empirical degree distribution to known models of distribution. When exploring 29 plant-pollinator and 24 plant-frugivore networks and examining their degree distribution separately for each

plant and animal species set, they found that most of the networks (65.6%) show a degree distribution of nodes that follows a decaying power-law. More precisely, the distribution is of the form $P(k) \propto k^{-\gamma} \exp(-k/k_x)$ in which k_x is a cut-off value. 22.2% show a scale invariance, best fitted by a power-law distribution ($P(k) \propto k^{-\gamma}$). The remaining networks either fit a fast decaying distribution, notably an exponential distribution ($P(k) \propto \exp(-\gamma k)$) or show no fit at all. Good fits for the power-law are found for intermediate values of the species degree (generally ≤ 30) while the decaying power-law is observed for higher values.

The authors not only described some features of mutualistic webs but they looked at plausible explanations of these features. Thus, the power-law fits may invoke an idea about the generalization level of the communities: specialists species are abundant in the community, in contrast to rare generalist species. The truncation of the power-law fit indicates that super generalists are rarer than expected: they are even rarer than in most non-ecological networks having the scale-free feature. While exploring their dataset further, they suggested that the observed exponential truncation may be the result of the existence of forbidden links between plants and animals: some species are constrained by their phenotypic traits, limiting the number of other species with which they can interact.

More recent studies dismissed some of the hypotheses proposed by Jordano et al. [2003]. Among others, Vázquez [2005] proposed an alternative in explaining the truncation of the power-law. Based on the assumption that the degree distribution of each species is a function of its frequency of interaction, they built a null model able to reproduce the empirical degree distribution of most of the studied communities. They showed that the truncation of the power-law can be explained even by the unique fact that the degree distribution and the frequency of interaction are correlated. Okuyama [2008], stated that the method used by Jordano et al. [2003] in the regression fitting may be unreliable. He proposed a maximum likelihood approach in finding the parameters for the fits, and found considerable disagreements in parameter values proposed by Jordano et al. [2003] and parameters given by his approach. Not only did he fit a power-law model and a truncated power-law model to the data, but he proposed a gamma distribution model as well. In many of the networks, the gamma distribution appeared to be the best fit.

To summarize, methods used by Jordano et al. [2003] show a predominance of a truncated power-law behaviour in the degree distribution indicating the abundance of specialists and the rarity of super generalists, while a maximum likelihood method shows that a gamma distribution is the best fit to the degree distribution. The degree distribution of some networks is also best fitted by a power-law or an exponential distribution.

2.3.2.2 Nestedness

Ecological mutualistic networks have been found to be highly nested. Nestedness was defined by Bascompte and Jordano [2007] to be a pattern of interaction in which specialists interact with species that form perfect subsets of the species with which generalists interact. After analysing 52 mutualistic networks, Bascompte et al. [2003] observed a highly nested structure in them. In fact, they estimated an index of nestedness for each network using the Nestedness Calculator Software [see Atmar and Patterson, 1993]. They then compared the nestedness index of the empirical networks with the nestedness index given by a null model composed of randomly assembled networks. 70% of the seed dispersal and 80% of the pollination networks have a nestedness index that departs significantly (at 5% level) from the null model. Furthermore, they investigated the relationship between the complexity of the web (indicated by its number of interactions) and the nestedness value and found that as the network complexity increases, they become more nested.
Other researchers investigated these nestedness patterns further. A nested structure was also observed by Dupont et al. [2003] in a pollination network in the high-altitude desert of Tenerife. Guimarães et al. [2006] confirmed the finding not only in pollination and seed dispersal networks but in ant-plant mutualistic networks as well. Even in interactions between sea anemones and their associated fish species, nestedness was observed [Ollerton and Allen, 2007]. Nestedness is consequently revealed to be a common feature in mutualistic networks. Nestedness in mutualistic networks has also been shown to exist regardless of the time and place data is sampled [Onielsen and Bascompte, 2007].
More recently, the measurement of nestedness has been the object of substantial controversy. Ulrich et al. [2009] argued that the metric used previously to measure nestedness is biased because the null model for the comparison is not constrained enough, leading to an exaggeration of the proportion of nested networks found in previous studies. Following this, Joppa et al. [2010] concluded that the use of a more constrained null model and other metrics in the characterization of nestedness tempered the high nestedness findings in ecological networks, but still give a dominant proportion of nested ecological networks.

2.3.2.3 Asymmetries

The nested structure implies another feature of mutualistic networks, namely their asymmetry. Asymmetry is also referred to disassortativity (see 2.1.1). In fact, nestedness gives an high asymmetric structure in terms of degree: rarely connected species (specialists) connect only with highly connected species (generalists), while generalists interact with both special-

ists and generalists. These asymmetries were observed by Vázquez and Aizen [2004].
Mutualistic networks also exhibit asymmetry in terms of strength. An interaction between two species can be weighted by the frequency of interaction. The strength of a species is then the sum of frequencies of interactions between the species and all its partners, e.g. the sum of the frequency of the visits a plant receives from pollinators. The concept of strength is related to the concept of dependency in the sense that the dependence of an animal on a plant is the proportion of the visits of the animal to this particular plant over all the visits given by the animal in the network. Bascompte et al. [2006] found a highly asymmetric structure of the dependencies: plant species depend strongly on animal species which depend weakly on the plant.

2.3.2.4 Modularity

A network exhibits a modular or a compartmentalized structure when it is organized into subgroups, namely modules, such that species within a module interact strongly with one another but share weak interactions with species in the other modules. Dicks et al. [2002] detected a strong evidence of compartmentalization in plant-insect mutualistic networks. They also observed that each module is composed of species in the same biological class or sharing some particular morphological traits.
Some time later, Olesen et al. [2007] worked further on the modularity characteristic by analysing a large dataset of pollination networks. They observed a modular pattern increasing with the number of species: large networks (>150 species) exhibit a modularity structure, while small networks (< 50 species) do not. As in the case of Dicks et al. [2002], they detected some convergent traits belonging to species in the same module and proposed an explication of the modular structure by the process of evolution.

2.3.3 Mutualistic Network Models

Once some insights concerning the structure of mutualistic networks are acquired, the possible process generating such features and the possible implications of such a structure on the network persistence remain two important questions for ecologists. One way to approach these questions is the construction of ecological network models able to generate the characteristics found empirically, and/or give theoretical predictions of the robustness of the networks.
In this context, an attempt to explain the truncation of the power-law observed in the node degree distribution was provided by Guimarães Jr et al. [2007]. They proposed a dynamical

network model based on the B-A model [Barabási and Albert, 1999], but modified in such a way that it generates a truncated power-law degree distribution instead of a power-law distribution. Modifications were made to the relative growth of the two mutualistic sets: either one set has a higher probability of acquiring a new node than the other one, meaning a difference in growth rate between the two groups, or one of the sets has a limiting size and stops growing. Both of those two mechanisms typify a significant difference in species richness between the two sets. This characteristic is also observed in the dataset used by Jordano et al. [2003] which shows the truncated power-law behaviour in the degree distribution. Thus, they demonstrated that the truncation of the power-law can emerge from processes associated with differences in species richness between the two sets.

In a similar work, Medan et al. [2007] built a model that can generate both the truncated power-law distribution and the nestedness features of mutualistic networks. Their self-organized network model is a dynamical model that keeps the total number of species in the two sets and the total number of interactions fixed. However, the connections are rewired during the process using the same principle as the preferential attachment rule: the allocation of the interaction to already well connected nodes is favoured. The rewiring operation is stopped depending on the match between the level of nestedness and the degree distribution of the modelled network and the empirical network. Since the model does not take the frequency of visits into account, the authors suggested that the correlation between the frequency of interactions and the degree distribution highlighted by Vázquez [2005] cannot be considered as a cause of the power-law truncation. They argued that the element responsible for the observed behaviour of mutualistic networks is primarily an elaborated preferential attachment rule, leading to both a truncated power-law and a high level of nestedness.

Subsequently, more elaborate models have been built. The bipartite cooperation model, able to replicate the degree distribution, the nestedness behaviour and the level of modularity was proposed by Saavedra et al. [2009]. The model construction is governed by two rules. The first is the specialization rule which determines the number of interacting partners a species can have, depending on a reward trait parameter drawn from a normal distribution. The second is the interaction rule which selects the most suitable species to be involved in the interaction, depending on the complementarity of traits between the pair. Their model was also found to be suitable for bipartite networks of manufacturer-contractor interactions, suggesting its relevance for a larger domain. By means of such a model, they demonstrated that the structure of interaction networks is strongly dependent on the complementarity in traits at the level of individual interactions.

So far, several models able to reproduce some features of mutualistic ecological networks,

such as the shape of the degree distribution, the nestedness pattern, and the modularity have been proposed. Thus, plausible explanations for the structure of ecological networks have emanated from those models. We can for instance list the existence of forbidden links, the correlation between the degree distribution and the frequency of visits, a high richness ratio between the two interacting sets, the preferential attachment rule and the level of complementarity of interaction between each pair of species. Simultaneously to those structural models, models built for future predictions were elaborated. We will focus on such models in the next section.

2.3.4 Robustness of Ecological Mutualistic Networks

The topological characteristics of mutualistic networks contribute enormously to their robustness, defined to be their resistance to species loss.
Jordano et al. [2006] presented a preliminary approach to the question after their discovery of the nestedness structure and the shape of the degree distribution. In a model, they simulated a perturbation by the removal of species in decreasing order of degree. Then, they observed the effects of such removals on the fraction of preserved links. Removal of only a small fraction of the most connected nodes had a dramatic effect on the network robustness. It also appears that the network is more sensitive to the loss of plant species than to the loss of animal species. Additionally, networks that exhibit a power-law degree distribution seem to be less robust than networks with a truncated power-law degree distribution. However, the nestedness pattern implies interactions among the most generalist plants and animals. Those interacting generalists form a dense core to which the rest of the web is linked. Thus, their removal may rapidly lead to a collapse of the entire network.
Successive removal of nodes is one way to simulate perturbation in an ecological network model in order to study the robustness of the network. Another way leading to the same goal is to simulate the loss of habitats. This approach was developed by Fortuna and Bascompte [2006]. Their model describes the changes in species abundance by means of the evolution of the fraction of available patches occupied by each species. The dynamical model has been tested for empirical networks, and for two different generated null model networks, namely a null model that does not reproduce the degree distribution and the nestedness, and a more conservative one which maintains the degree distribution approximately but not the nestedness. They found that as more habitats are destroyed, the rate of extinction increases. For each network, there is a threshold value of habitat loss such that beyond this value the communities suddenly start to collapse. Before the values of habitat destruction reach this threshold, real communities are seen to be less tolerant to habitat loss than sim-

ulated networks. The fraction of surviving species is smaller in real communities than in the simulated networks. Nevertheless, it appears that the threshold value is higher for the real communities than for the null model communities, meaning that the nested structure and the shape of the degree distribution in real communities allow a longer persistence. The proposed explanation for this robustness behaviour lies in the distribution of specialists and generalists. Since specialists rely on small number of species, they are the first to go extinct. However, real communities show a high number of specialists, explaining the earlier decay of real communities. Real communities persist longer because those many specialists share only few links with other species: their isolation does not cause an extinction cascade.

The effect of the nested structure of ecological networks was given particular attention in Bastolla et al. [2009]. The authors elaborated a mutualistic model defined by differential equations that describe the dynamics of species in the community. The model gives the number of species that can coexist in the community by incorporating the negative effect of the competition for resources between species in the same set and the positive effect of the mutualism between species in the two sets. Parameters of the models were first estimated from real networks which are known to be highly nested, and then from randomly simulated networks of which the bulk show no pattern of nestedness. Their results revealed that real communities show a higher increase in biodiversity (computed to be the total number of co-existing species) compared to randomly generated networks. The more nested the network is, the higher the number of species that can coexist in the network.

The asymmetric structure of mutualistic networks has also been found to have a major role in network robustness. In the study in Bascompte et al. [2006] detected a highly asymmetric structure in term of species dependences, they also observed the consequences of such structure for biodiversity maintenance. Recall that in mutualistic networks, plants have been observed to depend strongly on animal species while animal species depend weakly on plants. The authors defined an equilibrium state for the community such that the network is still in equilibrium when the product of mutual dependencies (of an animal on a plant and of the plant on the animal) is kept small. Two situations satisfy this requirement: either the plant and the animal depend weakly on each other, or if an animal (respectively a plant) depends strongly on a plant, the plant (respectively the animal) should depend weakly on the animal. Since real communities exhibit an asymmetry of mutual dependence, they meet the second criterion for stability. Thus, the asymmetric dependency in terms of strength has been revealed to be a positive factor for the coexistence of species. Furthermore, the authors observed a predominance of weak dependencies in real networks, which satisfies the first equilibrium criterion.

As with asymmetry in terms of dependencies, the effects of the observed asymmetry in terms

of degree in ecological networks have also been studied recently by Abramson et al. [2011]. The authors simulate the perturbation of a mutualistic network by the destruction of habitats. In order to make a comparison, they tested the model for different generated networks ranging from the most asymmetric to the least asymmetric, and for real pollination networks. Logically, specialist species should be more vulnerable to habitat destruction than generalist species. Surprisingly, this evidence was not found in pollination networks. The explanation lies in the degree asymmetry: it appears that in asymmetric networks, generalists and specialists are equally affected by habitat destruction, oppositely to the case of symmetric networks. In an assortative network, specialists interact with other specialists. After extinction of some specialists, their specialist partners also lose most of their interactions and become extinct easily. However, in an asymmetric network even if some specialists become extinct, the other specialists are preserved.
In summary, it seems that the non-random features of mutualistic ecological networks enhance their robustness. In a simulation of a perturbation where specialists are removed first, the abundance of specialists prevent a phenomenon of extinction cascade. Because of the nestedness feature implying asymmetric interactions in terms of degree, specialists and generalists are equally negatively affected by habitat destruction, in contrast to assortative networks. In addition, the asymmetry in dependencies keeps the networks stable. Nevertheless, mutualistic communities respond very badly to specific attacks such as those targeting the most connected species.

2.3.5 Phylogenetic Signal in Ecological Networks

As we have seen, several mechanisms have been proposed to explain the patterns governing the structure of ecological networks. Nevertheless, those explanations are mainly based on current ecological processes (such as preferential attachment or forbidden links due to non-matching traits), neglecting the fact that the observed structure may just be a result of the past evolutionary history of the species involved. Some studies have been performed in this direction. Those studies aim to detect a phylogenetic signal in the network, that is to say to detect a tendency of phylogenetically similar species to resemble each other.
Rezende et al. [2007a] were among the first to explore the role of phylogenetic history in the structure of mutualistic ecological networks. They showed that the nestedness pattern is partially explained by phylogeny and phenotypic complementarity. For this purpose, they created interaction matrices such that the existence of an interaction depends only on the phenotypic complementarity between the species pair. They started by considering a set of possible phenotypic traits, such as the corolla length of a pollinated flower, or the tongue

length of a pollinator. They then simulated the evolution of those phenotypic traits along phylogenetic trees under a Brownian motion model of evolution. The Brownian motion works as follows: starting at the root of the tree, a value is assigned for each trait. This trait value is then changed randomly following a normal distribution on every branch of the tree until the tips are reached. The variance of the normal distribution is also rescaled in proportion to each branch length. This process gives a value that can be assigned to each phenotypic trait of the species at the tips. The probability of each pairwise interaction was then defined according to the degree of matching between each pair of species. These probabilities were used to construct an interaction matrix. They performed this simulation using phylogenetic trees with contrasting degrees of hierarchy, ranging from the least hierarchical one, namely a star phylogeny, to the most structured one. For each species, a phylogenetic signal was measured to quantify how closely related species sharing the same traits tend to be. The degree of nestedness of the resulting networks was also analysed. They found that in contrast to phylogenies with a low degree of hierarchy, very hierarchical trees produce a high phylogenetic signal and highly nested networks. The degree of nestedness also increases with the number of phenotypic traits considered in the simulation. Consequently, they suggested that phenotypic diversity and phylogenetic history contribute to the nested structure of the networks. Furthermore, they examined a real bird-plant frugivory community with a known phylogeny. Some phenotypic traits of the species in the community were recorded. They then tested whether a phylogenetic signal can be detected in: (1) the phenotypic traits of the species, (2) in their degree (number of interactions of the species). The phylogenetic signal was considered significant when its value departs significantly (at 5% level) from a random permutation of the traits or the number of interactions among the species. They found that by contrast to the set of plants, both the phenotypic traits and the number of interactions tend to resemble in closely related bird species.

The same group explored further the detection of phylogenetic signals in real mutualistic communities [Rezende et al., 2007b]. They analysed a large number of pollination and frugivory networks of which phylogenies are known. Their goal was to observe the extent to which closely related species tend to share interaction patterns. For this purpose, they examined two components of interaction patterns: (1) the features governing the specialization level, namely species degree and its quantitative extension, species strength; (2) the partners identity of each species. When considering the data of species degree and species strength, the presence or the absence of phylogenetic signals was defined using statistical tools in comparative data analysis [Blomberg et al., 2003]. To detect whether closely related species tend to interact with the same partners, they began by constructing a phylogenetic distance matrix and an ecological distance matrix for each set of plants and animals. The ecological

distance matrix stores ecological similarity between species, which is defined by the proportion of the shared interactions between two species over the total number of species they interact with. A positive correlation between those two matrices indicates the presence of a phylogenetic signal. They found that for more than one-third of the analysed phylogenies (that is for 39% of the communities), phylogenetically related species tend to have a similar number of interactions. Only a small fraction of the communities show evidence of a phylogenetic signal when species strength is considered. For 42.7% of the phylogenies, a positive correlation was found between phylogenetic and ecological distance matrices, proving a tendency of phylogenetically related species to interact with the same partners.

A recent paper [Takemoto and Arita, 2010] proposed an evolving network model based on evolutionary processes to investigate the effect of the evolutionary history on the structure of mutualistic networks. The model is a growing network model starting from a small initial number of fully-interacting plants and animals, and governed by two processes. First, with probability p, a new plant is added to the network. This new plant is assumed to result from the mutation of a randomly selected pre-existing plant, Consequently, the new plant inherits phenotypic traits from its parent and has a high probability to interact with the animals the parent is linked to. However, the phenomenon of divergence reduces this probability to a probability q. Second, with probability $1 - p$, a pre-existing plant evolves and acquires new traits, allowing an interaction with a new added animal. The model parameters p and q are estimated from real data, knowing the total number of plants and animals and the total number of links in the real web. Next, they compared networks simulated under their model with real networks and networks simulated under the bipartite cooperation model described earlier [Saavedra et al., 2009]. Their model has been proved to be able to reproduce the nestedness pattern observed in real networks, similarly to the bipartite cooperation model. A suggested explanation for the nestedness lies in the mutation-divergence process: new plants, acquired from the mutation of the parents, interact with only a subset of the species the parent interacts with, because of divergence. Additionally, the model was shown to reproduce approximately the heterogeneous shape of the degree distribution of nodes. By reproducing those two well-known patterns using a simple model based on some mechanisms of evolution, their model suggests a non-negligible role of evolutionary history in the architecture of mutualistic networks.

So far, the existence of phylogenetic signals has been found to be in evidence in ecological networks. Furthermore, attempts to model the long evolutionary process resulting in current interaction webs reinforce the finding.

2.4 Conclusion

Complex interaction network studies are mainly motivated by the need to represent and explain the structure of real-world networks. Structural properties such as the small-world effect or the scale-free degree distribution have been observed to be common in most real-world networks. Consequently, most evolving network models aim to reproduce those properties by modelling the underlying possible mechanisms leading to such structure. In the case of ecological networks, more features are involved. If the study is limited to mutualistic interaction webs only, the skewness in the node degree distribution is one of their major properties. This skewness indicates generally either a power-law distribution, or more frequently a truncated power-law distribution and a gamma distribution. We can also list the nestedness pattern, leading to another ecological property which is the asymmetry in terms of degree. Quantitative studies also show an asymmetry in terms of strength and more recently, a modular pattern has been detected. Furthermore, knowing the global features of ecological networks helps in biodiversity preservation, in the sense that models based on the current observed structure allow prediction of the resistance of the networks to perturbations. It has for instance been proven that generalist species interacting with other generalists deserve particular attention in conservation since they form the mainstay of the network. Interest has also been focused on explaining the observed features of ecological networks. While some researchers proposed explanations relying on mechanistic processes such as the preferential attachment rule, the difference in size between the two interacting sets, or the correlation between the degree distribution and the strength distribution, other researchers focused on more biological explanations such as complementarity of phenotypic traits or external environmental factors causing barriers in the interactions. Additionally, the importance of evolutionary history in explaining those patterns has been demonstrated. In some real networks, it has been observed that phylogenetically related species tend to share the same patterns of interaction. However, these studies have been mainly concentrated on detection of phylogenetic signal. Few have explored models based on the process of evolution itself. This issue will be tackled in our model: motivated by the finding of a phylogenetic signal in real networks, our model will be based on the evolutionary history of the network and will be tested on the dataset used in Rezende et al. [2007b] in order to see the extent to which phylogenetic history can explain some network properties. Before getting to the model description, details on phylogenetic analysis tools, useful to understand correctly our model, will be discussed in the next chapter.

Chapter 3

Phylogenetic Analysis

According to Darwin's theory of evolution, all living organisms are related and have descended from a common ancestor. Darwin's theory prompted the idea of exploring the evolutionary history of a set of organisms, more explicitly the evolutionary relatedness among the organisms: this is the aim of phylogenetic analysis. Although Gregor Mendel, the father of modern genetics, published his works on heredity only a few years after Darwin's 'On the origin of species' [Darwin, 1859], the so-called 'Modern Synthesis' combining the theories of Darwinian evolution and Mendelian genetics, was only developed in the middle of the 20th century. After the discovery of the molecular structure of DNA, the molecule that carries genetic information from one generation to the next, in 1953, molecular methods have gained a real importance in phylogenetic analysis. Helped by tremendous advances in sequencing, the exploration of the history of evolution is nowadays mainly focused on molecular phylogenetic techniques.
Reconstructing the evolutionary history of a set of organisms involves an attempt to build a phylogenetic tree, and a model of evolution. In some cases, defining a model of evolution allows the estimation of model parameters such as the tree topology and branch lengths. Basic principles used in models of evolution will be discussed in the first section of this chapter. In the second section, we will talk about phylogenetic trees. The last section will focus on the description of the phylogenetic analysis software package used in our model.

3.1 Molecular Evolution Models

The evolution of DNA or amino acid sequences is usually modelled as a continuous time Markov process. Such a process, described by the variable $x(t)$ evolving through time t, is

generally characterized by:

- A set of the possible states of the variable x: $S = (S_1, \ldots, S_N)$
- Transition probabilities stored in a transition matrix $P(t)$. The rows and columns of $P(t)$ represent the states $S_1, \ldots, S_N$ and each element $P_{ij}(t)$ is the probability that state S_i changes to state S_j in time t. For the matrix elements to form a probability distribution, each matrix row should sum to 1.

A Markov process has the Markov 'memoryless' property: the conditional probability of being at a future state given the present state depends only on the current state and not on earlier states. Mathematically speaking, the 'memoryless' property means:

$$P\left(x(t_2) = S_i | x(t_1) = S_j, x(t_0) = S_k\right) = P\left(x(t_2) = S_i | x(t_1) = S_j\right) \qquad \text{for} \qquad t_0 \leq t_1 \leq t_2 \tag{3.1.1}$$

The evolution of a site in a sequence can be described by such a variable $x(t)$. Possible states of x are the nucleotide characters (A, C, G, T) for a DNA sequence and the 20 standard amino acids for an amino acid sequence.
The process is generally assumed to be stationary: the probability of substituting a character a at time t_0 by another character b at time t_1 depends only on the time interval $(t_1 - t_0)$. The stationarity property also ensures that the rate at which a state changes to another state is independent of the time.
For a short time dt, $P(dt)$ is approximately given by:

$$P(dt) \approx (I + Qdt) \tag{3.1.2}$$

Where I is the identity matrix and Q is known as the instantaneous rate matrix in which Q_{ij} is the rate at which state S_i is replaced by state S_j. The diagonal elements Q_{ii} of the rate matrix Q are defined such that the row sums are all zero.
The probability substitution matrix $P(t)$ is multiplicative, in the sense that $P(s)P(t) = P(s+t)$, for any time intervals s and t. This property is a result of the process being Markovian and stationary.
In fact, if we denote each element of the matrix $P(s)P(t)$ by $P_{ij}(s,t)$, we have:

$$P_{ij}(s,t) = \sum_k P_{ik}(s) P_{kj}(t) \tag{3.1.3}$$

$$= \sum_k \left[P\left(x(s) = k | x(0) = i\right) P\left(x(t) = j | x(0) = k\right)\right] \tag{3.1.4}$$

The stationary property allows us to write:

$$P\left(x(t)=j|x(0)=k\right)=P\left(x(t+s)=j|x(s)=k\right) \tag{3.1.5}$$

With equation 3.1.5, equation 3.1.4 becomes:

$$\begin{aligned}
P_{ij}(s,t) &= \sum_k \left[P\left(x(s)=k|x(0)=i\right)P\left(x(t+s)=j|x(s)=k\right)\right] \\
&= \sum_k \left[P\left(x(s)=k|x(0)=i\right)P\left(x(t+s)=j|x(s)=k,x(0)=i\right)\right] \qquad \text{(from equation 3.1.1)} \\
&= \sum_k \left[\frac{P\left(x(s)=k,x(0)=i\right)}{P(x(0)=i)}\frac{P\left(x(t+s)=j,x(s)=k,x(0)=i\right)}{P(x(s)=k,x(0)=i)}\right] \\
&= \sum_k \left[\frac{P\left(x(t+s)=j,x(s)=k,x(0)=i\right)}{P(x(0)=i)}\right] \\
&= \sum_k \left[P\left(x(t+s)=j,x(s)=k|x(0)=i\right)\right] \\
&= P\left(x(t+s)=j|x(0)=i\right) \\
&= P_{ij}(t+s)
\end{aligned}$$

Thus, with the multiplicativity property and equation 3.1.2, we get:

$$P(t+dt)\approx P(t)P(dt)\approx P(t)(I+Qdt)$$

In the limit of small dt, we get:

$$P'(t)=P(t)Q \tag{3.1.6}$$

Solving equation 3.1.6 gives

$$P(t)=e^{Qt}=\sum_{n=0}^{\infty}Q^n\frac{t^n}{n!}$$

The matrix exponential can be computed directly by a diagonalization of Q. In this case, we have:

$$Q=U^{-1}\Lambda U \qquad \text{and} \qquad e^{Qt}=U^{-1}e^{\Lambda t}U$$

where U is a matrix containing the eigenvectors of Q and Λ is a diagonal matrix composed of the eigenvalues of Q.
When the process runs for a very long time, it approaches an equilibrium state:

$$\lim_{t\to\infty}P_{ij}(t)=\pi_j$$

The equilibrium state is represented by the vector of equilibrium frequency $\pi=(\pi_1,\ldots,\pi_N)$ composed of the probabilities of being at each state S_i respectively. The elements of the

frequency vector should sum to 1.
Substitution models can also have the time reversibility property. A model is time reversible when it is invariant under the inversion of time scale, that is to say, if the evolutionary process is watched in reverse, it cannot be distinguished from when it is watched normally. A stationary reversible model satisfies the property:

$$\pi_i Q_{ij} = \pi_j Q_{ji} \qquad \forall i, j \in (1, \ldots, N)$$

- **Examples of Substitution Models**

One of the first DNA substitution models was proposed by Jukes and Cantor [1969]. This model makes simple assumptions such as an equal substitution rate for all nucleotides. Equilibrium frequencies are also all equal for all nucleotides ($\pi_A = \pi_C = \pi_G = \pi_T = \frac{1}{4}$). Thus, the rate matrix has the form:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{pmatrix} -3\alpha & \alpha & \alpha & \alpha \\ \alpha & -3\alpha & \alpha & \alpha \\ \alpha & \alpha & -3\alpha & \alpha \\ \alpha & \alpha & \alpha & -3\alpha \end{pmatrix} \end{array}$$

Assuming that all nucleotides are substituted at the same rate was found to be too simplistic. In fact, transitions (substitutions between nucleotides A $\leftrightarrow$ G and C $\leftrightarrow$ T) occur more frequently than transversions (substitutions between nucleotides A $\leftrightarrow$ G and C $\leftrightarrow$ T). Thus, the Kimura model [Kimura, 1980], with a transition and a transversion rate, was proposed. Like the Jukes & Cantor model, this model assumes equal frequencies of the nucleotide bases. The rate matrix of the Kimura model has the form:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{pmatrix} -2\beta - \alpha & \beta & \alpha & \beta \\ \alpha & -2\beta - \alpha & \beta & \alpha \\ \alpha & \beta & -2\beta - \alpha & \beta \\ \beta & \alpha & \beta & -2\beta - \alpha \end{pmatrix} \end{array}$$

Other more complicated DNA substitution models were proposed later. For instance, the F81 model [Felsenstein, 1981] was inspired by the Jukes-Cantor model with only one rate parameter, but the equilibrium frequencies are not assumed to be equal. In the HKY85 model [Hasegawa et al., 1985], the principles of the F81 and the Kimura model were combined: the rate of transversion and the rate of transition are distinguished and the frequencies are

not assumed equal. The GTR (general time reversible) model [Tavaré, 1986] generalizes all the models by assuming different rates and different frequencies for each nucleotide. Its instantaneous rate matrix is given by:

$$Q = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{pmatrix} . & \pi_C\alpha & \pi_G\beta & \pi_T\gamma \\ \pi_A\alpha & . & \pi_G\delta & \pi_T\epsilon \\ \pi_A\beta & \pi_C\delta & . & \pi_T\theta \\ \pi_A\gamma & \pi_C\epsilon & \pi_G\theta & . \end{pmatrix} \end{array}$$

The diagonal elements of the matrix are set such that each row sums to zero.
For a more detailed review of substitution models, see Liò and Goldman [1998].

Reconstructing the evolutionary history of a set of organisms cannot be achieved by the construction of an evolutionary model only: it also requires a phylogenetic tree, which we discuss in the next section.

3.2 Phylogenetic Trees

A phylogenetic tree is a schematic representation of the evolutionary process in which the current organisms form the tips of the tree, and are joined according to the similarities and differences in their genetic or physical features. Infering a phylogenetic tree is done by finding a tree that best describes the given data. Alignments of molecular sequences (DNA or amino acid sequences) are most often used as datasets on which to base this inference. Binary sequences composed of a presence (1) or an absence (0) can sometimes be used to describe some physical traits of the species.

3.2.1 Branch Lengths and Molecular Clock Hypothesis

An edge length in a phylogenetic tree represents the phylogenetic distance between two separated nodes. A phylogenetic distance is usually expressed in terms of expected number of substitutions per site. When a character within a sequence evolves into another character by a certain time interval t, the number of substitutions between these two characters can be known if the rate of substitution is given.
In some cases, the branch lengths represent chronological time. This is the case when they are estimated from fossil records of some of the species involved.

When building a phylogenetic tree, the molecular clock hypothesis is sometimes assumed. This hypothesis presumes that the rate of evolutionary change is constant over time, that is to say, the divergence of sequences is assumed to occur at the same constant rate at all points in the tree [Durbin et al., 1999a]. The molecular clock hypothesis is often modelled by an ultrametric phylogenetic tree: the sum of the lengths from the root down to a leaf node is invariant, regardless of the considered leaf node. If a molecular clock is assumed, the branch lengths given in terms of expected number of substitutions per site are proportional to chronological time.

3.2.2 Phylogenetic Tree Building

3.2.2.1 Methods based on pairwise distances

Some methods of phylogenetic tree building are based on phylogenetic distances. In the case of trees built from alignments of molecular sequences, a phylogenetic distance between a pair of sequences can be viewed as the expected number of substitutions per site that have occurred between the sequences and their common ancestor. This gives the branch lengths of the tree. Every substitution model defines its corresponding phylogenetic distance. For instance, the Jukes-Cantor model defines the phylogenetic distance to be $d = -\frac{3}{4}log(1 - \frac{4}{3}p)$ where p is the fraction of sites that differ between the two sequences.
One of the methods based on pairwise distances is UPGMA (Unweighted Pair Group Method with Arithmetic Mean), which is based on a sequence clustering process. It starts by considering each sequence as a cluster, then finds the two clusters with a minimum pairwise distance and combines them into a unique cluster. The combination will give a parent node placed at half-distance between the combined clusters. This process is repeated until only two clusters remain. The root of the tree is then placed between those two last clusters. The UPGMA method assumes the molecular clock hypothesis.
The Neighbour-joining method is also based on pairwise distances. The pairwise distance defined for the Neighbour-joining method is not only a function of the phylogenetic distance between the pair but also a function of the average distance between the considered pair and all other nodes. Firstly, the pair of closest nodes according to this new distance is found and linked to a common parent. Next, they are replaced by this new parent node and the distances that separate it from the other nodes are calculated, giving a new distance matrix. The process is repeated and the distance matrix updated at each time step until all nodes have been joined in this way.
Neighbour-joining and UPGMA produce binary trees, meaning that each edge splits into two

daughter edges. Indeed, all trees can be converted into a binary tree in the sense that nodes having more than two siblings are separated from the clade by the creation of an edge of length zero between them. However, unlike UPGMA, neighbour-joining creates an unrooted tree. In general, a real phylogenetic tree has a root. This root is the most common ancestor of the linked species. To root a tree, one can choose an outgroup: a group that is less related to the others than the others are between themselves. The root can be placed in the edge connecting this outgroup to the others.

3.2.2.2 Maximum Parsimony Method

Other tree building methods do not use pairwise distances. A commonly used one is the maximum parsimony method. This method aims to find the tree with the smallest number of evolutionary changes, namely with a minimum number of character substitutions. This method does not build the tree directly, but instead assigns a cost to all possible trees and chooses the one with a minimum cost. The assignment of the score is based on the number of evolutionary changes needed to produce the observed data. It can happen that several trees with the same total number of substitutions are selected as the most parsimonious ones. A maximum likelihood approach can also be applied in phylogenetic tree construction. Like the maximum parsimony method, it selects the best tree among all possible trees. We will look more closely at this method in section 3.2.3.

3.2.3 Probabilistic Approach

Tree inference can also be approached probabilistically. One commonly used method is maximum likelihood. Like the maximum parsimony method, the maximum likelihood method considers all possible trees and takes the one with the maximum likelihood. We are concerned about computing the likelihood of each tree with a particular topology and set of branch lengths. This is equivalent to computing the probability of the states of the n sequences at the tips of the tree T.

$$P(S_1, \ldots, S_n|T)$$

where the state of a sequence (S_i) is defined by the states of the sites composing the sequence. Computing the likelihood of the tree amounts to calculating the probability of each particular site u of the n sequences to be at the states $S_1^u, \ldots, S_n^u$ respectively: $P(S_1^u \ldots S_n^u|T)$. Then, the full likelihood of the sequences at the leaves of T is calculated by multiplying this probability at each site. T represents the topology and branch lengths of the tree.

$P(S_1^u \dots S_n^u|T)$ can be computed by Felsenstein's pruning algorithm [Felsenstein, 1981]. The objective of the algorithm is to compute the probability of all the leaves below the root node such that the state of the root node is fixed.
If we denote the root node by k and fix its state to a, this probability will be denoted by $P(L_k|a)$. This is given by the probability of all the leaves below its daughter nodes k_1 and k_2, namely $P(L_{k_1}|b)$ and $P(L_{k_2}|c)$, by the equation:

$$P(L_k|a) = \sum_{b,c} P(b|a,t_1)P(L_{k_1}|b)P(c|a,t_2)P(L_{k_2}|c) \tag{3.2.1}$$

where $P(b|a,t_1)$ is the probability of substituting a into b by the time t_1.
Iteratively, $P(L_{k_1}|b)$ and $P(L_{k_2}|c)$ are also computed using 3.2.1. The iteration ends when the considered nodes are leaves.
Then, one sums over all possible states of the root node to obtain the desired probability:

$$P(S_1^u \dots S_n^u|T) = \sum_a P(L_k|a)q_a$$

We can see that Felsenstein's algorithm requires the probability of substituting a character into another one. This is given by the model of evolution (see 3.1). In fact, once the evolutionary model is defined, it can be applied along each branch of the tree. The state S_j of a particular node is inherited from the state S_i of its parent node, according to the probability $P_{ij}(t)$ where the amount of time t is given by the branch length separating the node and its parent.
Alternatively to finding the right tree, the maximum likelihood approach can also be used to infer the model of evolution if the phylogenetic tree is given. This is equivalent to finding the values of the model parameters which maximize the likelihood:

$$\underset{\theta}{\operatorname{argmax}}\, P(S_1^u \dots S_n^u|T,\theta)$$

where θ denotes the set of model parameters.

Inference of phylogenetic trees and substitution models often requires a lot of computation regardless of the method used, especially when one has to deal with a large dataset or a large substitution matrix. To facilitate the analysis, several phylogenetic software packages are available. The one that will be used in our model construction will be described in the next section.

3.3 HyPhy

Hyphy, an acronym for Hypothesis testing using Phylogenies is a computational phylogenetic software package that performs a likelihood-based analysis on molecular evolutionary datasets [Kosakovsky Pond et al., 2005]. Whereas most phylogenetic software packages are mainly focused on phylogenetic tree construction, HyPhy emphasizes the study of the process underlying the evolution, such as the rates and the patterns of evolution.
The main application of HyPhy is to fit a phylogenetic tree with a model of evolution to a set of molecular sequences and give maximum likelihood estimates of the model parameters, such as the branch lengths of the tree and the substitution rates. For this purpose, HyPhy offers two different tools. On the one hand, a graphical user interface allows the user to perform the analyses easily. It provides access to several standard analyses commonly used in phylogenetic studies, such as a phylogenetic reconstruction or a model comparison. On the other hand, a batch language allows the construction of novel models and analyses. Using the batch language directly gives more control over the analysis one wishes to perform.

3.4 Conclusion

The reconstruction of the evolutionary history of a group of species is mainly based on finding an appropriate model of evolution and building a phylogenetic tree which best illustrates the history the species have in common. The evolutionary model describes the process of substituting sequence characters into each other within a specific time. The process is characterized probabilistically as a continuous time Markov process in which the transition probability is a function of a substitution rate matrix. Consequently, defining a model of evolution is focused on defining such a substitution rate matrix. A phylogenetic tree represents the evolutionary history schematically. Some phylogenetic tree constructions are based on distance matrices, such as UPGMA and neighbour-joining. Some of them consist of finding the best tree between all possible trees, either by taking the one with a minimum number of substitutions (maximum parsimony method), or taking the one that maximizes the likelihood. Maximum likelihood methods are adapted not only for a tree inference, but for any model parameter inference, such as the estimation of substitution rates. Phylogenetic analysis is the purpose of many software packages; the one used in this thesis is HyPhy, a phylogenetic software package focused on likelihood-based methods. In contrast to to several phylogenetic software packages, HyPhy concentrates mainly on the evolutionary process.
The model we present in chapter 4 is inspired by molecular models of evolution, and uses a

maximum likelihood approach to infer parameters. Details on model construction will follow in the next chapter.

Chapter 4

Model Construction

As highlighted in chapter 2, in mutualistic networks, phylogenetically related species tend to share the same patterns of interactions. The question of how much of the patterns of interactions are influenced by the phylogenetic history of the species remains to be investigated. By building an evolving network model that incorporates the process of co-evolution, we give an approach to that question. Our network model, suited for mutualistic networks, is based only on the phylogenetic history of the species. This chapter will be devoted to describing the method we used to construct the model. In the first section, we give a general overview of the model and by presenting its main objective. The second section provides a formal description of the model. Finally, we describe some experiments conducted under the model.

4.1 Model Overview

4.1.1 Mutualistic Networks

Our evolutionary network model is suited for mutualistic ecological networks, where two sets of species interact. Generally, we have a set of plants on one side, and a set of animals on the other side.
Such a mutualistic network is often represented by a bipartite graph, where the nodes (representing the interacting species) can be classified into two sets. Interactions in a bipartite graph exist between two entities belonging to two different sets, but not between entities within the same set. Figure 4.1 illustrates a mutualistic network represented by such a bipartite graph.

Another way to represent an interaction network is its adjacency matrix A, in which mem-

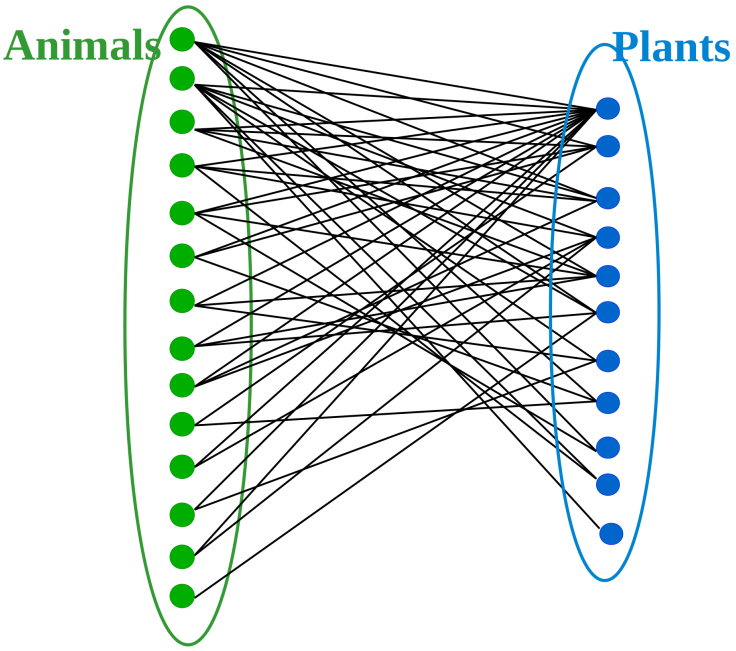

Figure 4.1: A seed dispersal interaction network between communities of birds and berries in Great Britain in 1988

bers of a set are placed on the rows and the members of the other set on the columns. The elements of A are defined such that:

$$a_{ij} = \begin{cases} 1 & \text{if entity } i \text{ and entity } j \text{ interact} \\ 0 & \text{otherwise} \end{cases}$$

An example of such an adjacency matrix is given in figure 4.2.

4.1.2 Purpose of the Model

Our model is an evolving network model. Based on the interaction network between two sets of species at the present time, and the evolutionary history of the species, the model gives a reconstruction of the potential interactions between species ancestral to the present species at any time in the past (see figure 4.3). The evolutionary history of the species is given by phylogenetic trees of the two sets. We consider only rooted phylogenetic trees in which branch lengths are directly proportional to the predicted evolutionary time between organisms. The phylogenetic trees are also assumed to be ultrametric (see 3.2.1).

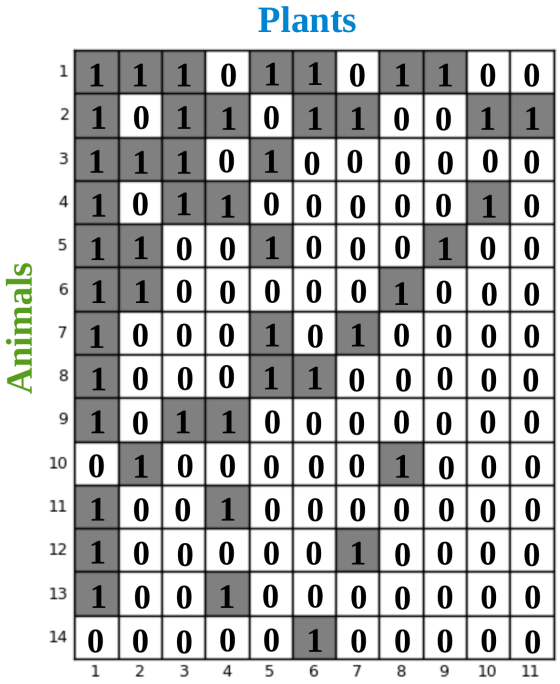

Figure 4.2: Adjacency matrix associated with the network in figure 4.1

4.2 Model Formalization

Since our main goal is to model the evolution of the link between each pair of species in the evolutionary history, the task is less complicated if we have a representation of all the species pairs present at any time. This will be done by combining the two phylogenetic trees into a unique tree.

4.2.1 Combined Tree

Points on a phylogenetic tree represent species that are ancestral to those on the leaf nodes. We built the combined tree such that each point on it is associated with a pair of contemporaneous points, one from each phylogenetic tree, and represents the potential interaction between the ancestral species on those trees.
The method for the combination is illustrated in 4.4. We start by constructing the oldest branch of the combined tree (Br 1A). This will be the combination of the pair of oldest branches of the two phylogenetic trees (Br 1 and Br A respectively). The branch Br 1A leads to a node which corresponds to the interaction between the most recent common an-

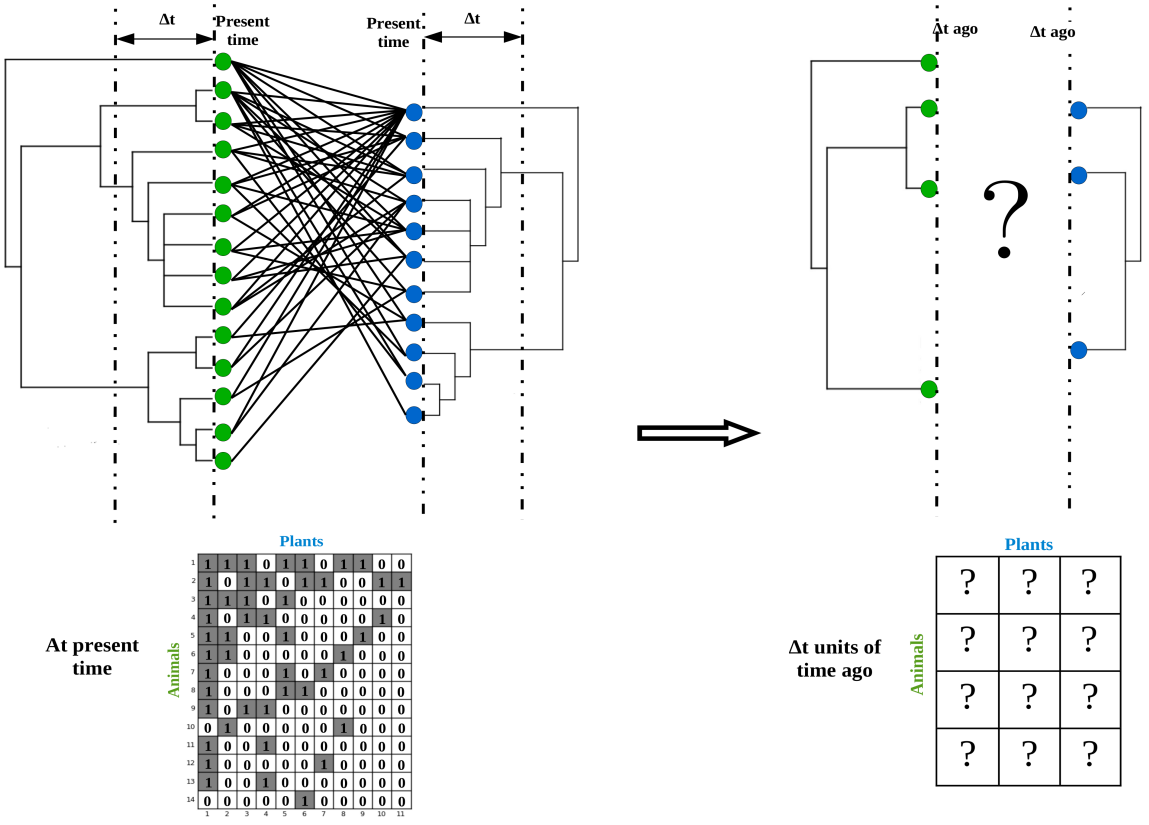

Figure 4.3: The model aims to reconstruct the network at any time in the past

cestors (root node 1 and root node 2) of the two sets of species. Next, the split at Node 1 in phylogenetic tree 1 is represented in the combined tree by the split of the branch Br 1A into Br 2A and Br 3A. By going forward in time, we split Br 2A and Br 3A simultaneously into Br 2B, Br 2C and Br 3A, Br 3C. Those simultaneous splits correspond to the split at Node A in Phylogenetic tree 2. We continue following this process to create new branches in the combined tree, until we reach leaf nodes, representing species at the present time (Node 2, Node 3, Node B and Node C).

Each leaf node in the combined tree then represents the interaction between a pair of species, one from each side of the network at present time. For instance, the node ending Br 2B represents the interaction between Node 2 and Node B. Thus if we know the interaction network at present time, we know the state of each leaf node in the combined tree.

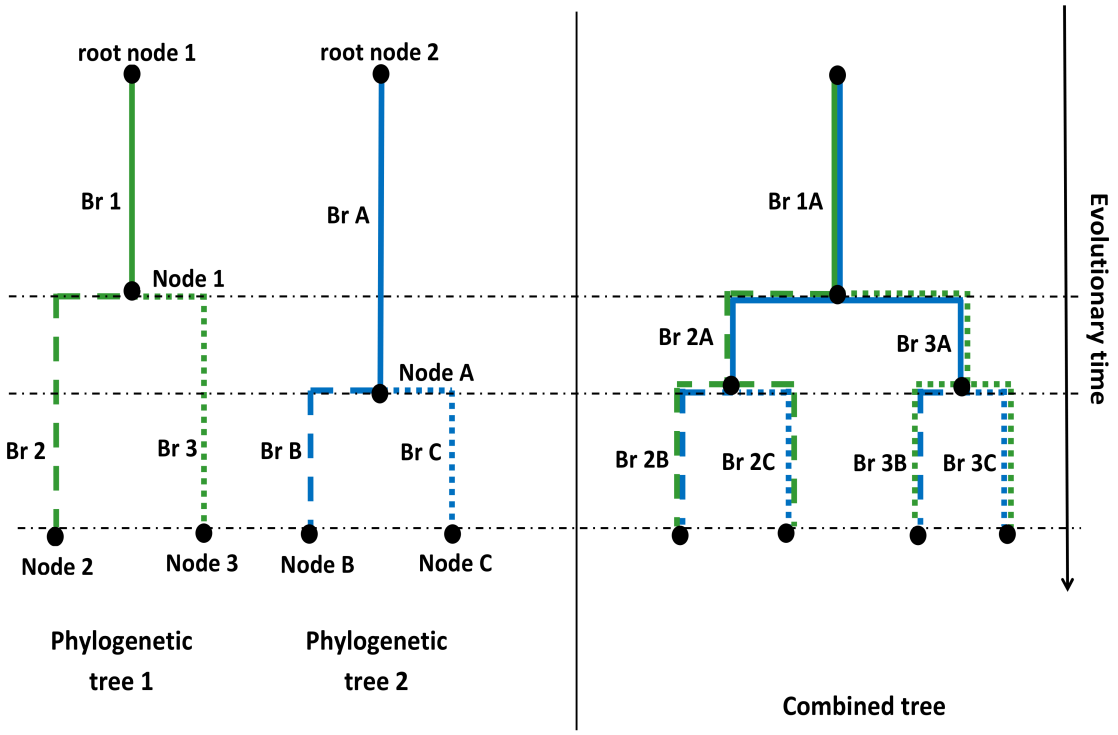

Figure 4.4: Example of a combination of two phylogenetic trees

4.2.2 Evolution of an Interaction along each Branch of the Combined Tree

With the combined tree representing the pair of species present at any time, the evolution through time of the interaction between each species pair can now be modelled. We model this evolution as a continuous time Markov process (see 3.1), and consider the evolution of an interaction similarly to the evolution of a sequence character in models of molecular evolution.

The set of states is $S = (s_0, s_1)$ in which s_0 refers to an absence of interaction and s_1 refers to a presence of interaction. The transition probability $P(t)$ is a 2×2 matrix, such that $P_{01}(t)$ is the probability of going from state s_0 to state s_1 within an amount of time t. The diagonal elements of $P(t)$ ($P_{00}(t)$ and $P_{11}(t)$) represent the probability of retaining the presence or absence of an interaction.

We assume a stationary reversible model: the rate of gaining or losing an interaction does not change over time and the model is invariant under the direction of time. Let us denote by μ the gain-loss rate parameter. The rate matrix Q is in our case a 2×2 matrix. Since

a stationary reversible model should satisfy $\pi_i Q_{ij} = \pi_j Q_{ji}$ (for $i, j \in \{0, 1\}$), Q is therefore given by:

$$Q = \begin{array}{c} \\ s_0 \\ s_1 \end{array} \begin{array}{c} \begin{array}{cc} s_0 & s_1 \end{array} \\ \begin{pmatrix} -\mu\pi_1 & \mu\pi_1 \\ \mu\pi_0 & -\mu\pi_0 \end{pmatrix} \end{array} \tag{4.2.1}$$

where $\pi_0 + \pi_1 = 1$; π_0 and π_1 are the equilibrium frequencies of s_0 and s_1 respectively.
$Q_{01} = \mu\pi_1$ is the rate of gaining an interaction. Since in a rate matrix, each row is mathematically required to sum to zero, the rate at which the absence of an interaction is retained (Q_{00}) is set to $-\mu\pi_1$.
Consequently, our model has the gain-loss rate parameter μ and the equilibrium frequency vector $\pi = (\pi_0, \pi_1)$ as main parameters.
There is also a possibility to deal with a non-reversible model. In that case, one instead considers two rate parameters: the gain and loss rate parameters (μ_1 and μ_2 respectively) and the rate matrix has the form:

$$Q = \begin{array}{c} \\ s_0 \\ s_1 \end{array} \begin{array}{c} \begin{array}{cc} s_0 & s_1 \end{array} \\ \begin{pmatrix} -\mu_1\pi_1 & \mu_1\pi_1 \\ \mu_2\pi_0 & -\mu_2\pi_0 \end{pmatrix} \end{array}$$

Since results we got with a non-reversible model do not differ much from those obtained from a reversible model, we stuck to the simpler one, that is the reversible model.

4.2.3 Evolution of the Entire Network through Time

The evolution of the network through time is modelled by applying equation 4.2.1 to all the branches of the combined tree. We assign to each node of the tree a variable describing the presence or absence of an interaction. The state s_j of an interaction at a particular node n_1 is inherited from the state s_i of the interaction at its parent node n_0 ($\forall i, j \in \{0, 1\}$). This inheritance depends on the transition probability $P_{ij}(t)$ where t is given by the length of the branch separating n_0 and n_1. The state of the interaction at the root node is estimated from the observed probability distribution of interaction presence and absence in the networks.

4.2.4 Implementation

By representing the state of a node in the combined tree (that is, the presence or absence of an interaction) as a binary sequence of length one, we transform our problem into a special

case of the sequence evolution problem discussed in chapter 3. This allows us to make use of software developed for that problem: we took advantage of this by implementing our model in the HyPhy Batch Language (see 3.3), which allows modelling of binary sequence data. Our code takes as inputs the two phylogenetic trees of the two groups of species, and the adjacency matrix corresponding to the network at present time. HyPhy reads phylogenetic trees given by their standard Newick representation. Once the model is defined, HyPhy also allows us to perform simulations under the model in order to generate simulated networks. We are grateful to Prof. Sergei Kosakovsky Pond, one of HyPhy developers, for his contribution to the code implementation. He implemented not only the part of the code that constructs the combined tree but the part that deals with ancestral state reconstruction as well. His fruitful suggestions also helped us to improve the whole implementation, namely in the analysis part (model construction and parameter inference) and in the simulation part. The implementation can be accessed at `http://users.aims.ac.za/~ony/codes/full_model.bf`

4.2.5 Model of Branch Lengths

Although phylogenetic trees were available for each set of species in the interaction network, the branch lengths were typically unknown. As we do not expect the interaction networks to be informative for inference of the branch lengths, we resorted to a simplified branch length model.

We consider two forms of phylogenetic trees: internal branches are shorter than external branches, that is, branch lengths increase as we go from the root to the leaf nodes; and the opposite case of decreasing branch lengths from the root to the leave nodes.

We introduced two parameters to model branch lengths: λ which controls the variation of the branch lengths, and K which controls the relative scaling between the two trees. Each branch of a phylogenetic tree is obtained according to the following equation:

$$L_i = (\frac{d_i}{D})^{\lambda-1}$$

where L_i is the length of the branch leading to node i, d_i is the depth of node i (length of the path - see 2.1.1 - separating node i to the root) and D is the height of the tree.

Notice that $\lambda \in \mathbb{N}$ and when $\lambda = 1$, all branch lengths are equal.

When $\lambda > 1$, the branch lengths increase as we go from the root to the leaves of the tree.

When $\lambda < 1$, L_i is a decreasing function of d_i. We deal with values of λ given by 2^k for k in the range $(-5, -4, \ldots, 4, 5)$. Beyond that range, branch lengths become numerically indistinguishable from zero.

Multiplying the branch lengths in one of the two trees by K allows us to scale the two phylogenetic trees relative to each other. We consider the values of K in the range $(0.25, 0.5, \ldots, 3.25, 3.5)$. Beyond that range, we obtain overly long branch lengths, which leads to saturation (the state of the descendant node is no longer a function of the state of the parent node).
We stuck to this model of branch lengths after consideration of a model of constant branch lengths (all branch lengths fixed to 1). Using this model of branch lengths noticeably improves the results of the model compared to using a model of constant branch lengths. Even if real branch lengths are known for one side of the communities (for instance for animals only), using them still leads to poor results compared to using our branch lengths model.

4.2.6 Inference of Parameters

We use a maximum likelihood approach (see 3.2.3) to infer the model parameters μ, π, λ, K. That is, we take the values of μ, π, λ, K that maximize the probability of observing the state of the interactions at present time. We first iterated over all possible combinations of values of λ and K. Each combination corresponds to an assignment of specific lengths to the branches of two phylogenetic trees. After construction of the tree combining those phylogenetic trees with fixed topology and branch lengths, the model of evolution is defined and a maximum likelihood analysis performed using the model and the corresponding combined tree. Corresponding optimized values of μ and π are given by each analysis. The specific maximum likelihood analysis (that is, for a specific combination (λ, K)) which shows a maximum value of the likelihood function is then selected.
We can also construct the evolutionary history of the interaction network by finding the ancestral states in the combined tree that maximize the joint likelihood function for the inferred parameter values. That is, by maximizing the following expression over $S_1, \ldots, S_N$:

$$P(S_1, \ldots, S_N | \mu, \pi, \lambda, K)$$

Here, $S_1, \ldots, S_N$ represent the states of the interactions between all contemporaneous pairs of ancestors of the present species.

4.3 Experimentation

Once we built the model, we tested it using empirical networks.

4.3.1 Data

We validated the model using the interaction networks used in Rezende et al. [2007b]. These are derived from 53 communities composed of 31 plant-pollinator and 22 plant-frugivore networks, spanning a large geographic range. While the dataset of interaction networks are available in the paper itself, we obtained the phylogenies of the species by request from the authors. The phylogenetic trees are rooted and ultrametric trees. Branch lengths, given in units of expected number of substitutions per site, were available for only 18 animal communities. Since those available branch lengths belong to only one side of the bipartite graph (the animals), as pointed out earlier (see 4.2.5), we decided to explore further the phylogenies to which we assigned branch lengths generated by our model.

4.3.2 Analysis

We analysed the 53 empirical networks. For each of the networks and for all the values of λ and K, we inferred the values of μ and π, and noted down the given optimized likelihood of each analysis.
For a maximum likelihood estimation, we retain only the particular values of λ and K maximizing the likelihood of the analysis. That gives us unique values of μ,π,λ and K for each network.

4.3.3 Simulations

For a comparison with existing models of degree distribution, we are also interested in optimizing the likelihood of the degree distribution rather than of the full data set. To do this, we still consider all the values of K and λ, and get for each of those values the inferred parameters μ and π. Then, for each pair of λ and K, we generated 1000 simulations under the corresponding μ and π for each network. In this case, we will only optimize λ and K by maximizing the likelihood of the degree distribution.
For each empirical network, apart from the set of simulations under our model, we also performed a set of 1000 simulations under a random model (see 4.3.5). Notice that both the random simulated matrices and the matrices simulated under our model have the same size as the corresponding empirical matrix.

4.3.4 Ecological Network Properties Study

Study of some ecological network features in both the empirical networks and the simulated networks gives an indication of how far the structure of ecological networks is influenced by the past history of the species involved. We focussed on exploring two main ecological network properties: the distribution of the node degree, and the nestedness feature.

4.3.4.1 Models of Node Degree Distribution

We obtain an estimate of a model of probability distribution of the node degree (see 2.1.1) from the networks simulated under our model. Considering at the same time all 1000 simulated networks, the probability $P(k)$ of obtaining a particular node degree k is given by the fraction of the number of times k is observed over all observed values of node degrees. In some cases, some values of k are observed in the empirical network but not in the simulated networks. Since simply setting $P(k)$ to be equal to 0 for those values would harshly imply a rejection of our model, we instead use pseudocounts: for each simulated network, we added one extra count to those values of k and to all values of k observed in the simulated networks [Durbin et al., 1999b].
We then used this model of node degree distribution to calculate a likelihood value of the corresponding empirical data. For each network, we retained the values of K and λ that maximize this likelihood.
As highlighted in 2.3.2.1, some other well established degree distribution models are usually used to fit the node degree distribution in ecological networks. We also considered those probability distribution models, namely a power-law model, an exponentially truncated power-law model, an exponential model, and a negative binomial model. We fitted those models to the empirical networks using the maximum likelihood approach.

Power-law Distribution

The probability distribution of a discrete power-law over an integer variable k is of the form:

$$p(k) = Ck^{-\alpha} \qquad \forall k \geq k_{min}$$

where C is the normalizing constant, k_{min} is the minimum possible value of k, and α is the scaling parameter [Seal, 1952]. We assume that $\alpha > 1$ since distributions with $\alpha \leq 1$ are not normalizable.
In the case of a discrete power-law in which $k_{min} = 1$, by solving $\sum_{k=k_{min}}^{\infty} p(k) = 1$, the

normalizing constant C is given by:

$$C = \frac{1}{\zeta(\alpha)}$$

where

$$\zeta(\alpha) = \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \qquad \text{is the Riemann zeta function.}$$

The logarithm of the likelihood function is given by:

$$L = \log \prod_{i=1}^{n} \frac{k_i^{-\alpha}}{\zeta(\alpha)} = -n \log \zeta(\alpha) - \alpha \sum_{i=1}^{n} \log k_i \tag{4.3.1}$$

for each degree of node k_i observed in the dataset.
An estimate of the best power-law fit to the data is then given by taking the value of the scaling parameter α that maximizes L. This maximum likelihood estimate of α is the solution of:

$$\frac{\partial L}{\partial \alpha} = -n \frac{\zeta'(\alpha)}{\zeta(\alpha)} - \sum_{i=1}^{n} \log k_i = 0$$

For each of the 53 empirical networks, we estimated such a power-law degree distribution model, and computed the value of the logarithm of the likelihood of the empirical data, using equation 4.3.1.

Truncated Power-law Distribution

The probability distribution function of an exponentially truncated power-law is given by:

$$p(k) = Ck^{-\alpha} e^{\frac{-k}{\lambda}} \qquad \forall k \geq k_{min}$$

where α is the scaling parameter and λ the truncation parameter.
The truncation parameter λ can be interpreted as the cut-off value, or the value of k at which the distribution departs from a power-law distribution [Jordano et al., 2003]. It is then assumed that $\lambda \geq k_{min}$
When $k_{min} = 1$, the normalizing constant C is given by:

$$C = \frac{1}{\lambda^{(1-\alpha)} \Gamma(1-\alpha, \frac{1}{\lambda})}$$

[Clauset et al., 2007] where $\Gamma()$ is the upper incomplete gamma function, and is given by:

$$\Gamma(1-\alpha, \frac{1}{\lambda}) = \int_{\frac{1}{\lambda}}^{\infty} k^{-\alpha} e^{-k} dk$$

Since the exponentially truncated power-law distribution is a continuous distribution, we discretized it in order to fit our discrete data. Hence, we set the probability of a node degree k being j to be:

$$\begin{aligned}
P(k=j) &= \int_{j-0.5}^{j+0.5} p(k)dk \\
&= \frac{1}{\lambda^{(1-\alpha)}\Gamma(1-\alpha,\frac{1}{\lambda})} \int_{j-0.5}^{j+0.5} k^{-\alpha} e^{\frac{-k}{\lambda}} dk \\
&= \frac{1}{\lambda^{(1-\alpha)}\Gamma(1-\alpha,\frac{1}{\lambda})} \lambda^{(1-\alpha)} \left[\Gamma(1-\alpha,\frac{j-0.5}{\lambda}) - \Gamma(1-\alpha,\frac{j+0.5}{\lambda})\right] \\
&= \frac{\Gamma(1-\alpha,\frac{j-0.5}{\lambda}) - \Gamma(1-\alpha,\frac{j+0.5}{\lambda})}{\lambda^{(1-\alpha)}\Gamma(1-\alpha,\frac{1}{\lambda})}
\end{aligned}$$

In the special case where $j = k_{min} = 1$, the value of $P(k = k_{min})$ is obtained by integrating over the range $[k_{min}, k_{min} + 0.5]$ only.
The logarithm of the likelihood function is then given by:

$$L = \sum_{i=1}^{n} \log \frac{\Gamma(1-\alpha,\frac{k_i-0.5}{\lambda}) - \Gamma(1-\alpha,\frac{k_i+0.5}{\lambda})}{\lambda^{(1-\alpha)}\Gamma(1-\alpha,\frac{1}{\lambda})} \tag{4.3.2}$$

We estimated the values of the parameters α and λ by maximizing the likelihood function L, and noted the likelihood value of the empirical data given this truncated power-law model.

Exponential Distribution

The probability density function of an exponential distribution is given by:

$$p(k) = Ce^{-\lambda k} \qquad \forall k \geq k_{min} \tag{4.3.3}$$

where $\lambda > 0$ is known as the rate parameter. For $k_{min} = 1$, the normalizing constant C is $C = \lambda e^{\lambda}$.
Given a dataset $(k_1, \ldots k_n)$, the logarithm of the likelihood function of an exponential distribution is:

$$L = \prod_{i=1}^{n} \log p(k_i) = \sum_{i=1}^{n} [\log \lambda + \lambda - \lambda k_i] = n \log \lambda + n\lambda - \lambda \sum_{i=1}^{n} k_i$$

The maximum likelihood estimate of the rate parameter λ was obtained by solving:

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} + n - \sum_{i=1}^{n} k_i = 0$$

which gives the estimate

$$\widehat{\lambda} = \frac{1}{\overline{k} - 1} \qquad \text{with } \overline{k} = \frac{\sum_{i=1}^{n} k_i}{n}$$

As with the case of the truncated power-law distribution, the value of the likelihood of the empirical data was computed by discretizing the density function given in 4.3.3. Consequently, the probability of observing a node degree $k = j$ was set:

$$\begin{aligned} P(k = j) &= \int_{j-0.5}^{j+0.5} p(k)dk \\ &= \lambda e^{\lambda} \int_{j-0.5}^{j+0.5} e^{-\lambda k} dk \\ &= e^{\lambda} \left[e^{-\lambda(j-0.5)} - e^{-\lambda(j+0.5)} \right] \end{aligned}$$

so that the log likelihood of the empirical degree distribution is given by:

$$\sum_{i=1}^{n} \left[\log \left(e^{-\lambda(k_i - 0.5)} - e^{-\lambda(k_i + 0.5)} \right) \right] + n\lambda$$

Negative Binomial Distribution

The probability mass function of a standard negative binomial distribution over a non-negative integer k is given by:

$$f(k) = \binom{k + r - 1}{r - 1} p^r (1 - p)^k \qquad \forall k \in \{0, 1, 2, 3, \dots\}$$

where r is treated as a positive discrete parameter, $p \in [0, 1]$, and $\binom{A}{B} = \frac{A!(A-B)!}{B!}$ [Aragón et al., 1992].

Since in our case, the minimum possible value of k is 1, we instead consider a zero-truncated negative binomial distribution, of which the probability mass function, that we denote $g(k)$, has the same form as $f(k)$ up to a constant C.

$$g(k) = Cf(k) \qquad \Rightarrow \qquad \sum_{k=1}^{\infty} g(k) = C \sum_{k=1}^{\infty} f(k)$$

Since $g(k)$ is a probability mass function, we have $\sum_{k=1}^{\infty} g(k) = 1$.

We then have:

$$C \sum_{k=1}^{\infty} f(k) = 1 \qquad \Rightarrow \qquad C \left[\sum_{k=0}^{\infty} f(k) - f(0) \right] = 1$$

$$\Rightarrow \qquad C \left[1 - p^r \right] = 1 \qquad \Rightarrow \qquad C = \frac{1}{1 - p^r}$$

Consequently, the probability mass function of a zero-truncated negative binomial distribution is given by:

$$g(k) = \frac{1}{1-p^r}\binom{k+r-1}{r-1}p^r(1-p)^k \qquad \forall k \in \{1,2,3,\dots\}$$

A negative binomial distribution is well defined, and maximum likelihood estimates of r and p exist and are unique only when the sample variance exceeds the sample mean [see Aragón et al., 1992]:

$$\frac{\sum_{i=1}^n k_i^2}{n} > \frac{\sum_{i=1}^n k_i}{n}$$

For some of the empirical networks to which we want to fit a model of degree distribution, this criterion was not satisfied. We did not consider a negative binomial model for those cases.

The maximum likelihood estimates of p and r, and the value of the log likelihood of the empirical data was obtained using the following likelihood function:

$$L = -\log(1-p^r) + \log(\Gamma(k+r)) + \log(\Gamma(k+1)) - \log(\Gamma(r)) + r\log p + k\log(1-p)$$

where $\Gamma(n)$ is the Gamma function, and if n is a positive integer, $\Gamma(n) = (n-1)!$.

Since $r \in \{0,1,2,\dots\}$, a direct estimation of r from the likelihood function L turned out to be impossible because an infinite number of values had to be considered as r is not bounded in a certain range. We instead estimate its inverse $z = \frac{1}{r}$ which is bounded in the range $(0,1]$.

Notice that all the parameters of those degree distribution models were estimated numerically. That is, we implemented the above defined equations in a python script. The script can be accessed at `http://users.aims.ac.za/~ony/codes/degree_dis_fit.py`

4.3.4.2 Nestedness

Another property of empirical and modelled networks we investigated is the nestedness behaviour.

We measured the degree of nestedness using the *Aninhado* software [Guimarães and Guimarães, 2006] which does this by computing the NODF (Nested metric based on Overlap and Decreasing Fill) [Almeida Neto et al., 2008]. Recall that a network is nested when specialists

interact only with subsets of the species generalists interact with (see 2.3.2.2). When in the adjacency matrix, rows and columns are arranged from the most generalist to the most specialist, perfect nestedness will imply the absence of interaction between any species a_2, more specialist than a species a_1, and any species p_1 given that p_1 does not interact with a_1 (see figure 4.5). Consequently, a nestedness metric should take into account how much full each column and each row is, and whether presence of interactions in emptier columns and rows overlap with those in fuller columns and rows.

Aninhado first considers a measure of nestedness for each pair of rows and each pair of columns. Assuming that a pair of rows r_1 and r_2 (respectively, columns c_1 and c_2) are such that r_1 (resp. c_1) is located above r_2 (resp. to the left of c_2), *Aninhado* penalizes a situation in which r_2 is fuller than r_1 (resp. c_2 is fuller than c_1). For non-penalized pairs, *Aninhado* then computes the nestedness degree to be the percentage of filled cells in r_2 (resp. c_2) which overlap with filled cells in r_1 (resp. c_1). The NODF value of the entire matrix is then given by averaging the nestedness values of all the possible pairs of columns and rows. This value, ranging from 0 to 100, increases with the nestedness of the matrix.

We computed, for each network, the NODF value of the empirical network as well as those of 1000 simulations under our model. This allows us to report a p-value for each modelled network.

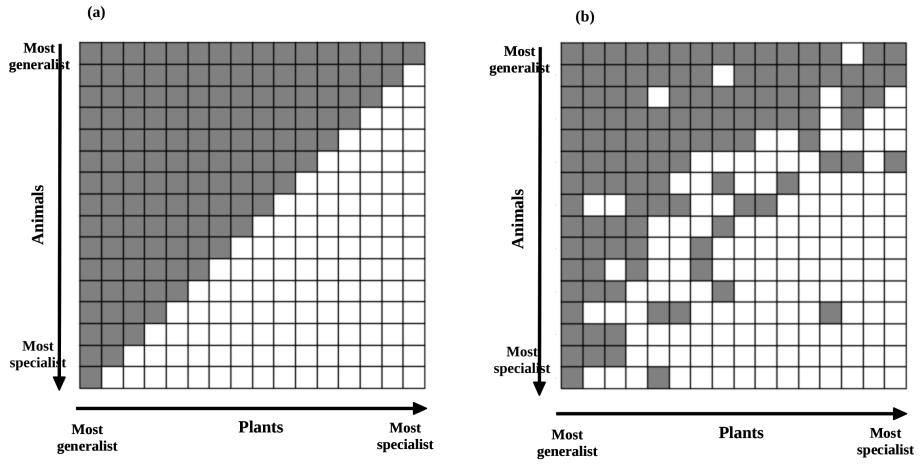

Figure 4.5: (a) A perfectly nested matrix with NODF value=100. (b) A frugivory network in Southern Spanish Mediterranean shrublands; cells filled in grey represent a presence of interaction; species are arranged from the most generalist to the most specialist; NODF value=79.75

4.3.5 Null Models

Null models are pattern-generating models that deliberately exclude a mechanism of interest, and allow for randomization tests of ecological and biogeographic data [Gotelli, 2001]. To statistically measure the significance of the shape of the degree distribution, and the nestedness behaviour in empirical and simulated networks, we generated two different null models.

Null Model for Degree Distribution Study

For each network, we generated a set of 1000 random matrices such that the presence of an interaction in a cell of a matrix is assigned with probability given by:

$$p = \frac{\#presence}{N}$$

in which $\#presence$ is the number of 1s in the corresponding empirical matrix and N is the size of the empirical network.
The null model of node degree distribution was obtained by estimating a model of degree distribution from those randomly generated matrices. We estimated this degree distribution the same way we estimated a degree distribution model from the matrices simulated under our model (see 4.3.4.1).
This is the Null model I in Bascompte et al. [2003], and is known to maintain the observed number of interactions probabilistically, while the interactions are reshuffled among all animal-plant pairs. The distribution of the node degree is then not conserved in this null model.

Null Model for Nestedness Study

For the nestedness study, we randomly generated 1000 matrices in such a way that for each of their elements a_{ij}, the existence of an interaction between species i and species j has probability p given by:

$$p = \frac{1}{2}\left(\frac{N_i}{C} + \frac{N_j}{R}\right)$$

where N_i is the number of 1s in row i, C is the number of columns; N_j is the number of 1s in column j and R is the number of rows.
This is Null model II in Bascompte et al. [2003]. The total observed number of interactions

is maintained probabilistically in each generated matrix, and the presence of an interaction is proportional to the level of generalization of both the interacting plant and the animal. In that way, each species conserves approximately its number of interactions. Consequently, this null model reproduces approximately the shape of the degree distribution, but lacks the nestedness behaviour.
By making use of those null model matrices, we also computed the relative nestedness value of each empirical and simulated matrix. This is given by:

$$N^* = \frac{N - \overline{N}_R}{\overline{N}_R}$$

where N is the value of the nestedness of the matrix (also referred as absolute nestedness) and $\overline{N}_R$ is the average nestedness value of the 1000 random replicates in the null model. The relative nestedness allows us to compare between networks, regardless of the size of the networks we wish to compare.

4.4 Conclusion

So far, we modelled the evolution through time of an ecological network of interactions. The model describes networks in which interacting species are classified into two interacting groups, namely bipartite networks. One can consider the examples of pollination or frugivory networks where species interact mutually between themselves. The phylogenetic history of the species give information on the evolutionary time by means of the branch lengths of the phylogenetic trees. The model reconstructs the potential interactions between ancestral species at any time in the past. We started the model construction by combining the two phylogenetic trees into a unique tree in order to have a representation of all pairs of species present at any time. We then used a probabilistic approach - a continuous time Markov process - to model the evolution through time, namely along each branch of the combined tree, of an interaction between any pair of species. A gain-loss rate parameter and a parameter for the equilibrium frequencies were introduced in the Markov process. Those parameters were inferred using a maximum likelihood approach. To test the model, we used 53 real-world pollination and frugivory networks. After inferring the parameters from those empirical networks, we simulated under our model and got the simulated networks of which we studied characteristics. We focused our attention mainly on two features: the shape of the degree distribution, and the nestedness behaviour of the networks. For the degree distribution study, apart from our model, we also considered other models: a power-law, a truncated power-law, an exponential, and a negative binomial model. To quantify the

importance of the studied properties, null models have been generated for both the degree distribution and the nestedness analyses.
While this chapter details the process we followed to construct the model, we will devote the next chapter to detail the outcomes of the model.

Chapter 5

Model Outcomes

This chapter is concerned with showing the results we got by simulating ecological networks under our model. Those results will be compared to the empirical networks and to the networks simulated under null models. As described in the previous chapter, those results will mainly concern the study of some features of both the empirical and simulated networks. The first section focuses on evaluating the likelihood values obtained by the inference and comparing them to likelihood values obtained using randomly generated phylogenies. The second section will be devoted to results obtained by studying the node degree distribution, while the third section will be dedicated to results concerning nestedness. We give the maximum likelihood estimates of the model paramters in section 4. In the last section, we will discuss the results by trying to answer how much of ecological network structures are influenced by the evolutionary history of the species. We will also attempt to clarify why some networks are well fit by our model and others are not, by taking into account assumptions we made in the model.

5.1 Likelihood Comparison

The likelihood of a model for a given data set can give information on how well the data are represented by the model. Since we want to know the importance of the evolutionary history of the species in our model, we looked at how likely it is to observe the current interaction network given the phylogenies we used.
For this purpose, we repeated the analysis but this time using randomly generated phylogenies. We generated random phylogenetic trees, under the constraints of being ultrametric and having the same number of leaves as the number of leaves observed in the corresponding

true phylogenetic tree. The generation of such a random phylogeny is described in Algorithm 1.

For each network, we generated 100 different random phylogenies and noted down the likelihood of the reconstruction. We expect to obtain higher likelihoods by using the real phylogenies instead of using the random ones.

Algorithm 1 Generation of a random phylogeny

```
input Desired number of leaf nodes: n
initialization Add the first node (root node)
repeat
  Generate a length l from an exponential distribution
  for all leaf nodes do
    Choose randomly between splitting the node or not
    if Splitting then
      Generate the two daughter nodes
      Assign a length l to the branches leading to the daughter nodes
    else
      Extend the branch leading to the node by length l
    end if
  end for
until Number of leaf nodes = n
```

Results and Interpretations

For 47 of the 53 networks, that is to say, for 88.68% of them, using the correct phylogeny is significantly better (at 5% level) than using a random phylogeny. Figure 5.1 shows the histograms of the likelihoods of 20 different network analyses. Further details concerning each network can be seen in table A.1. We can see that in most of the cases (17 networks out of 20), the likelihood of the analysis using the real phylogenies is higher than using randomly generated ones.

These results show the relevance of using the right phylogenies in constructing our model. In addition, they confirm that the structure of the current networks is partially influenced by the past evolutionary history of the species involved.

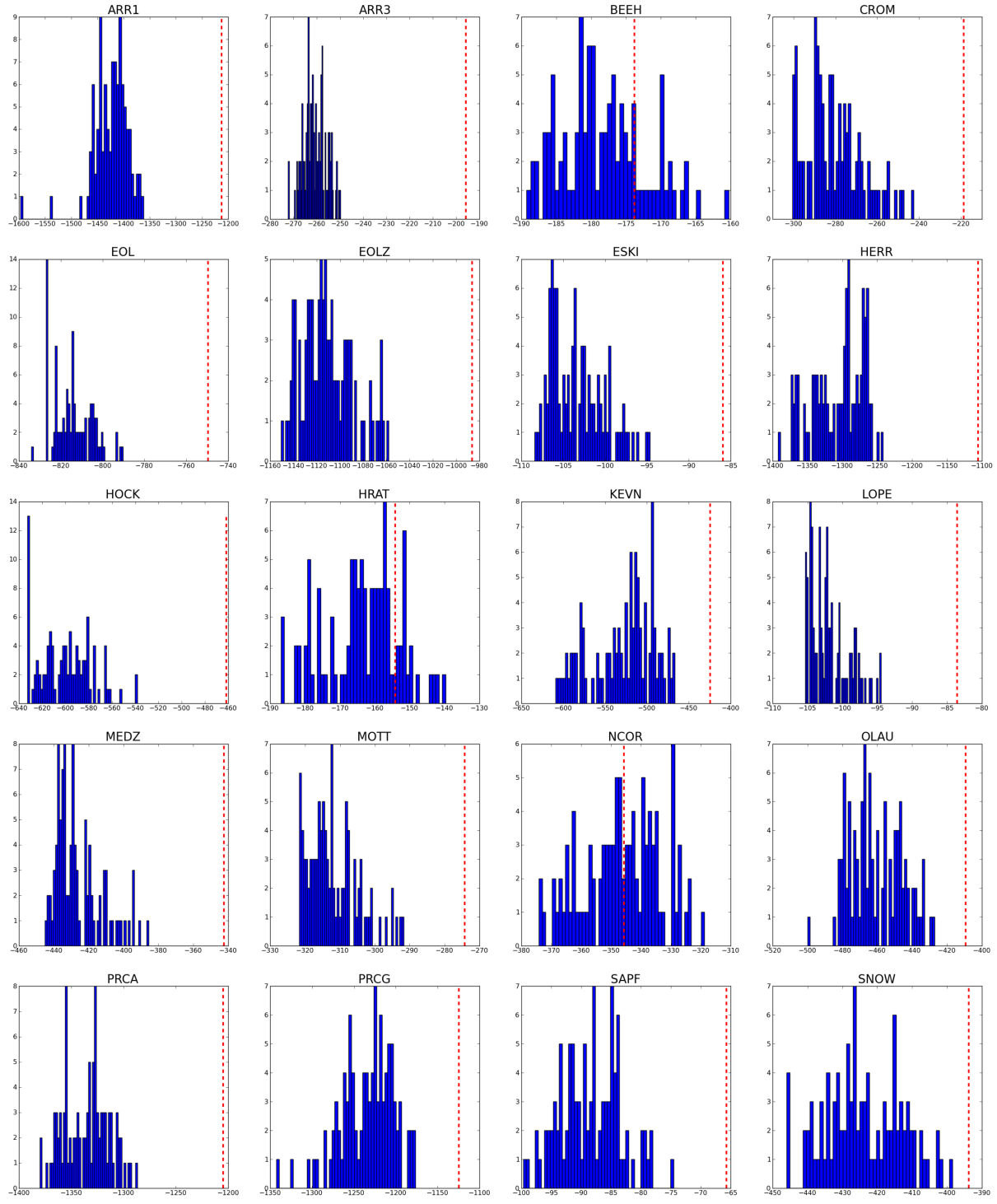

Figure 5.1: Likelihood histogram of the analysis of different networks. In each network, likelihood values when using random phylogenies are represented by the blue bars; the red broken line indicates the value of the likelihood when using the real phylogenies

5.2 Node Degree Distribution

5.2.1 AICc

To have a measure of the goodness of fit of the candidate degree distribution models (our degree distribution model, the power-law, the truncated power-law, the exponential, the negative binomial and the null models), we computed their AICc (Akaike Information Criterion corrected) scores [Burnham and Anderson, 2004], given by:

$$AICc = 2k - 2ln(L) + \frac{2k(k+1)}{n-k-1}$$

where L is the maximized value of the likelihood for the specific model, k is the number of parameters in the model and n is the size of the dataset. In our case, n is the number of species in the network.
$AICc$ decreases with the likelihood value L, but increases with the number of parameters k. The model with the lowest $AICc$ score is selected as the best among all models (see Table 5.1).
$AICc$ is generally used as an indication of the relative goodness of fit for several models. Individual $AICc$ scores are not informative about the acceptance or the rejection of a specific model. Consequently, we instead computed the $\Delta AICc$ value for each model and for each network:

$$\Delta AICc = AICc - AICc_{min}$$

where $AICc_{min}$ is the minimum $AICc$ score among all models for each specific network.
The best model for a network has $\Delta AICc = 0$. Models with $\Delta AICc \leq 2$ are considered to have substantial support. Models with $4 \leq \Delta AICc \leq 7$ have less support, and those with $\Delta AICc > 10$ are not supported [Burnham and Anderson, 2004].
On the one hand, we considered for each network the node degree distribution for animals and plants together. On the other hand, we also considered the node degree distribution for plants and animals separately for each network (see Figure 5.2). $\Delta AICc$ results are summarized in tables 5.2, 5.3 and 5.4.

5.2.2 Results

When the node degrees for animals and plants are considered in the same distribution, the $\Delta AICc$ results show us that 33.96% of the networks are best fit by our model (18 out of 53 networks). In addition, 2 other networks cannot reject our model.

35.85% (19 out of 73 networks) of the node degree distributions follow an exponentially truncated power-law behaviour, with another 8 networks for which the truncated power-law model cannot be discarded.
However, when the node degree distributions are fitted separately for plants and animals, fewer networks (16.04% of the communities, that is 18.87% of the animal communities and 13.2% of the plant communities) show evidence for our model. The predominant pattern of the truncated power-law fit (24.53% of the communities) confirms the node degree distribution studies by Jordano et al. [2003].

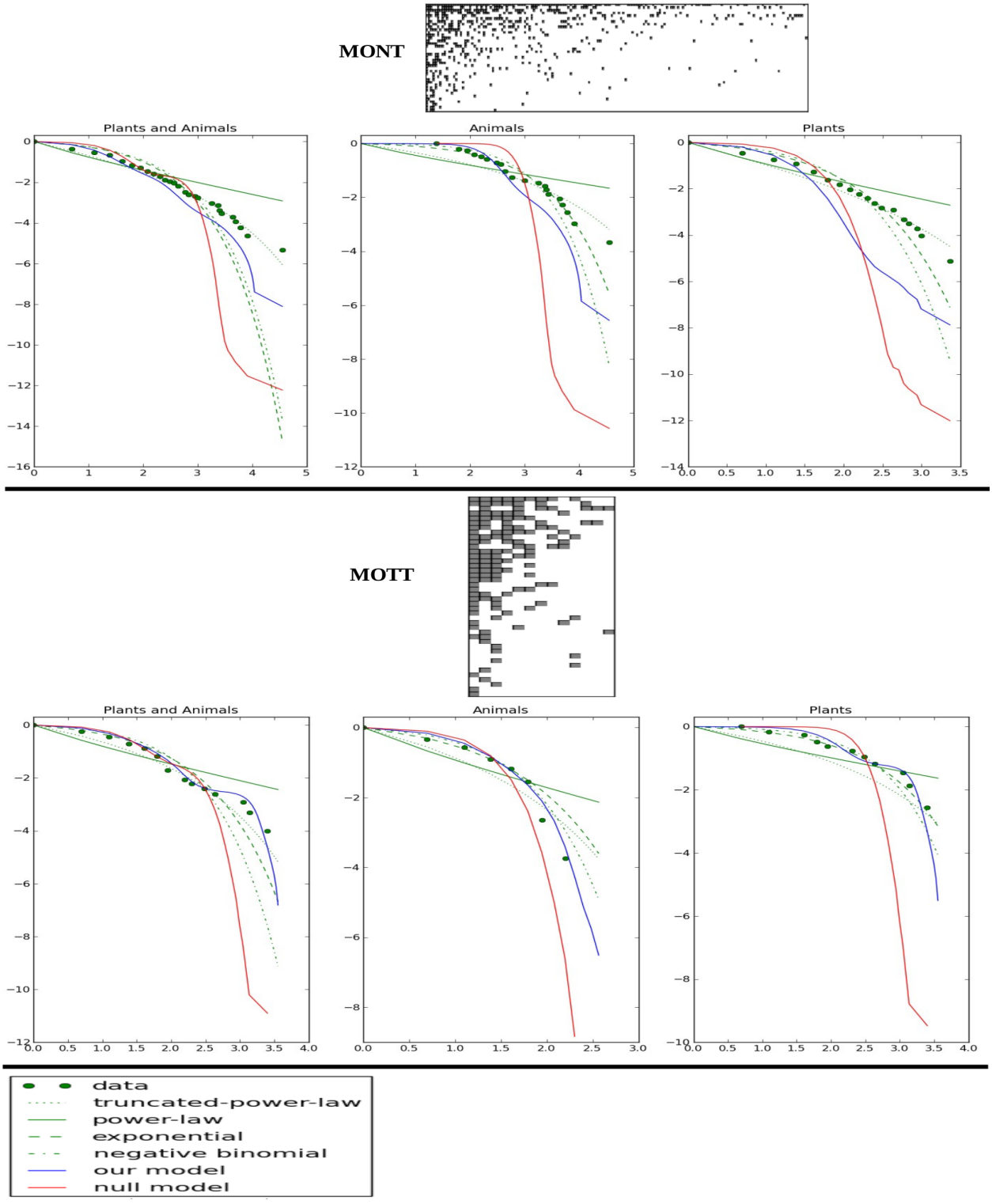

Figure 5.2: Log-log plot of the complementary cumulative distribution of the connectivities of examples of a frugivory network (MOTT) and a pollination network (MONT); when plants and animals are fitted together, the degree distribution in the frugivory network is best fitted by a truncated power-law and the degree distribution in the pollination network by our model.

Table 5.1: AICc scores of each model of degree distribution (for both animals and plants)

Networks	Our model	PL	TPL	Exp.	Neg.Bin.	Null model	Best fit
ARR1	865.007	830.163	807.335	871.25	853.073	1036.631	Truncated PL
ARR2	434.059	452.093	425.555	439.707	432.723	481.966	Truncated PL
ARR3	181.275	185.176	178.469	191.391	183.302	200.59	Truncated PL
BAHE	337.15	324.325	331.721	453.546	407.247	355.475	Power-law
BAIR	125.38	130.845	126.414	128.889	127.912	127.502	Our model
BEEH	226.861	245.292	226.873	224.486	223.287	252.317	Negative Bin
CACG	162.995	171.396	165.35	170.63	168.737	193.044	Our model
CACI	216.955	227.847	218.027	241.169	240.222	293.572	Our model
CACO	135.876	131.561	130.593	142.185	137.29	149.483	Truncated PL
CAFR	142.56	141.006	139.158	148.446	144.1	163.953	Truncated PL
CLLO	1753.815	1595.443	1603.389	1867.751	1733.285	2085.366	Power-law
CROM	276.512	300.979	287.808	359.257	340.228	305.564	Our model
DIHI	333.201	316.072	311.464	358.933	336.999	344.964	Truncated PL
DISH	207.982	209.07	208.473	223.973	215.421	235.678	Our model
DUPO	223.851	248.052	223.863	224.101	225.856	228.385	Our model
EOL	550.251	530.253	550.627	642.197	581.563	609.394	Power-law
EOLZ	647.091	733.346	655.43	642.727	650.641	752.68	Exponential
ESKI	118.545	121.179	109.395	109.675	109.406	118.21	Truncated PL
FROS	130.82	181.109	158.712	140.691	131.341	135.754	Our model
GEN1	97.776	99.086	102.646	105.412	100.375	105.239	Our model
GEN2	312.738	324.31	298.587	301.765	297.409	363.706	Negative Bin
HAMM	243.061	324.339	283.908	256.165	244.69	256.592	Our model
HERR	755.623	740.157	756.204	930.228	846.39	804.579	Power-law
HOCK	373.205	394.962	376.507	434.95	431.677	425.685	Our model
HRAT	186.797	239.826	206.811	185.284	178.415	240.315	Negative Bin
INPK	595.298	586.177	563.92	602.98	594.818	696.484	Truncated PL
KANT	152.125	193.604	165.89	163.248	171.554	154.953	Our model
KEVN	371.665	372.813	373.738	445.211	411.008	392.523	Our model
LOPE	124.529	162.731	141.1	129.685	125.721	124.015	random model
MACK	184.679	167.981	173.368	203.467	184.21	200.071	Power-law
MED1	214.73	213.982	208.0	240.711	235.278	234.324	Truncated PL
MED2	281.244	294.913	292.034	360.709	347.602	310.302	Our model
MEMM	355.083	378.088	344.857	350.695	352.805	427.934	Truncated PL
MOMA	103.705	102.252	100.912	107.445	102.552	109.678	Truncated PL
MONT	1149.772	1152.089	1114.076	1203.33	1150.204	1294.994	Truncated PL
MOTT	282.126	307.8	282.178	287.155	289.276	293.275	Our model
MULL	458.655	395.332	422.297	595.829	519.519	507.375	Power-law
NCOR	317.565	335.343	303.845	306.045	308.297	388.068	Truncated PL
NNOG	261.062	283.02	252.523	246.545	246.81	289.256	Exponential
OLAU	349.847	331.944	335.431	386.046	356.541	443.318	Power-law
PERC	356.799	359.637	348.833	382.004	371.885	394.884	Truncated PL
PRAP	281.702	295.059	282.931	300.042	292.609	285.282	Our model
PRCA	808.043	794.479	778.816	856.33	806.104	862.128	Truncated PL
PRCG	763.954	738.389	732.013	796.536	745.118	843.374	Truncated PL
RABR	263.113	294.204	266.607	276.404	269.127	274.581	Our model
RMRZ	385.236	397.009	371.408	388.379	383.006	433.371	Truncated PL
SAPF	106.296	103.74	101.979	102.965	101.042	110.515	Negative Bin
SCHM	146.791	152.327	153.445	176.541	165.121	159.1	Our model
SMAL	245.553	287.627	254.93	244.719	244.618	247.147	Negative Bin
SMRA	630.422	610.29	607.562	732.376	679.485	702.333	Truncated PL
SNOW	356.646	427.212	366.386	354.037	365.171	381.339	Exponential
WES	1545.265	1466.521	1408.433	1499.514	1438.177	1748.576	Truncated PL
WYTH	107.63	136.662	113.783	107.997	108.479	107.731	Our model

Table 5.2: Repartition of the networks according to the ΔAICc scores of each degree distribution model for animals and plants

ΔAICc / Models	0	≤2	≤4	≤8	≤16	≤32	≤64	≤128	≤256	≤512
Our Model	18	2	5	5	10	8	3	0	2	0
Power-law	7	6	2	6	7	13	9	3	0	0
Truncated PL	19	8	6	7	6	6	1	0	0	0
Exponential	3	6	2	8	9	4	8	9	3	1
Neg. Binomial	5	7	3	11	4	11	6	5	1	0
Null Model	1	1	4	3	7	12	8	12	3	2

Table 5.3: Repartition of the networks according to the ΔAICc scores of each degree distribution model for animals

ΔAICc / Models	0	≤2	≤4	≤8	≤16	≤32	≤64	≤128	≤256	≤512
Our Model	10	6	8	9	10	2	4	2	2	0
Power-law	12	4	3	10	5	12	7	0	0	0
Truncated PL	16	7	9	7	5	8	1	0	0	0
Exponential	6	3	5	7	12	9	6	4	1	0
Neg. Binomial	6	7	7	8	9	5	5	1	0	0
Null Model	3	0	2	4	13	9	9	7	4	2

Table 5.4: Repartition of the networks according to the ΔAICc scores of each degree distribution model for plants

ΔAICc / Models	0	≤2	≤4	≤8	≤16	≤32	≤64	≤128	≤256	≤512
Our Model	7	3	9	8	10	8	5	0	3	0
Power-law	7	2	5	4	9	14	11	1	0	0
Truncated PL	10	11	6	6	14	5	1	0	0	0
Exponential	16	7	3	11	7	6	2	1	0	0
Neg. Binomial	9	15	9	9	5	2	1	1	0	0
Null Model	4	5	3	2	6	10	11	8	2	2

5.2.3 Interpretation

By using a simple model based only on the phylogenetic history of the species, we obtain a better quantification of degree distribution than existing models in about 34% of the cases. This indicates the non-negligible influence of the evolutionary process in ecological network structures.
We also investigated whether the influence of the evolutionary history is more accentuated in some networks than in other ones, depending on some properties of the networks. For this purpose, we considered the networks according to their natures, their sizes and their nestedness values.

According to their natures

Given the following repartition of the networks regarding their natures and whether they are well fit by our model or not:

$$\begin{array}{c} \\ \text{pollination} \\ \text{frugivory} \end{array} \begin{array}{c} \begin{array}{cc} \text{our model} & \text{other models} \end{array} \\ \begin{pmatrix} 9 & 22 \\ 9 & 13 \end{pmatrix} \end{array} \tag{5.2.1}$$

we observed that although frugivory networks only represent 37.14% (13 out of 35) of the networks best fit by other models, 50% of the networks best fit by our model are frugivory networks. This suggests that compared to other models, our model may be more suitable for frugivory network rather than pollination network. One explanation may lie in the fact that animal communities in pollination networks we studied are composed of different insect species (bees, flies, beetles) while animal communities in frugivory networks are composed of species of birds only (and mammals only for one network). This is to say that animal species are phylogenetically more distant in pollination networks than they are in frugivory networks.
However, we tested the veracity of the assumption by performing a Fisher's exact test. It consists of calculating the conditional probability of getting the matrix 5.2.1 such that the marginal columns and rows sums are fixed. The significance of the result is quantified by a p-value, which is the probability of getting a matrix as extreme as or more extreme than the observed one. In our case, $p = 0.394$ which indicates a non significant association between network type and whether the network is well fit by our model.

According to their sizes

We observed that networks best fit by our model are much smaller (with an average number of species of 53.44) than networks best fit by other models (average number of species: 99.63). The significance of the difference was then tested by performing a T-test for two independent samples. The sizes of the networks are classified into two categories: networks best fit by our model, and those best fit by other models. The T-test gives an indication of the separateness of the two groups, and consists of checking whether the means of the two groups are statistically different from each other. We got a p-value of $p = 0.014$ which indicates that we can accept a significant difference (at 5% level) in size between the two categories.

According to their nestedness

Networks best fit by our model are observed to be more nested (with a mean of the nestedness values: 42.99) than networks best fit by other models (mean of the nestedness values: 33.43). Using a T-test for two independent samples, we obtained a p-value of 0.09. Thus the effect has only borderline significance and is rejected at the 5% level.

5.3 Nestedness Values

Computation of the degree of nestedness of each empirical network, of the corresponding 1000 networks simulated under our model, and of the 1000 networks simulated under the null model allowed us to report p-values. If a hypothesis, usually named the null hypothesis, is assumed, a p-value is the probability of observing by chance a test statistic at least as extreme as the one that is actually observed. A result is usually considered significant when the p-value is ≤ 0.05. In that case, when a significant result is observed, the null hypothesis can be rejected.
By considering two different such null hypotheses, two tests were performed:

- Firstly, we tested whether the null model for nestedness can be rejected. In this case, we assumed that the null model for nestedness could be a representation of the empirically observed degrees of nestedness (this is the null hypothesis). Networks from which we obtained a significant result (p-value ≤ 0.05) are the ones for which the null hypothesis is rejected, that is to say, for which the null model for nestedness is rejected.

- Secondly, we tested whether our model of nestedness can be rejected. The null hypothesis is in this second case that our model could represent the empirically observed degrees of nestedness. The p-value would be the probability of observing a nestedness value in the networks simulated under our model at least as extreme as the one observed empirically. Our model is rejected when this p-value is ≤ 0.05.

5.3.1 Results according to p-values

Details of the nestedness scores for empirical and simulated networks are given in table 5.5. For 43 of the 53 networks (81.13%), the hypothesis of a random behaviour (given by the null model) of nestedness in empirical networks can be significantly (at a p-value ≤ 0.05) rejected. It confirms the highly nested structure of mutualistic networks highlighted in Bascompte et al. [2003].
In addition, 24 of the 53 networks discard our model, which means that for the remaining 54.72% of the networks, the hypothesis that our model could be a representation of the empirically observed degree of nestedness cannot be rejected.
When scores of relative nestedness (see 4.3.5) are used instead of the absolute ones, we observed that all the networks rejecting our model regarding absolute nestedness values also reject our model regarding relative nestedness values. This case is also observed for networks rejecting the null model. Thus we got similar proportion of model rejection as when absolute nestedness were considered (81.13% reject the null model and 45.28% reject our model). Relative nestedness scores can be seen in table 5.6.
While the average values of nestedness of simulated networks are plotted against nestedness of empirical networks (see figure 5.3), we can observed that our model performs better for higher degrees of nestedness, compared to the null model. In fact, as the degree of nestedness increases, the regression line that fits the simulated data is closer to the line x=y than the line fitting the random networks is to the line x=y.

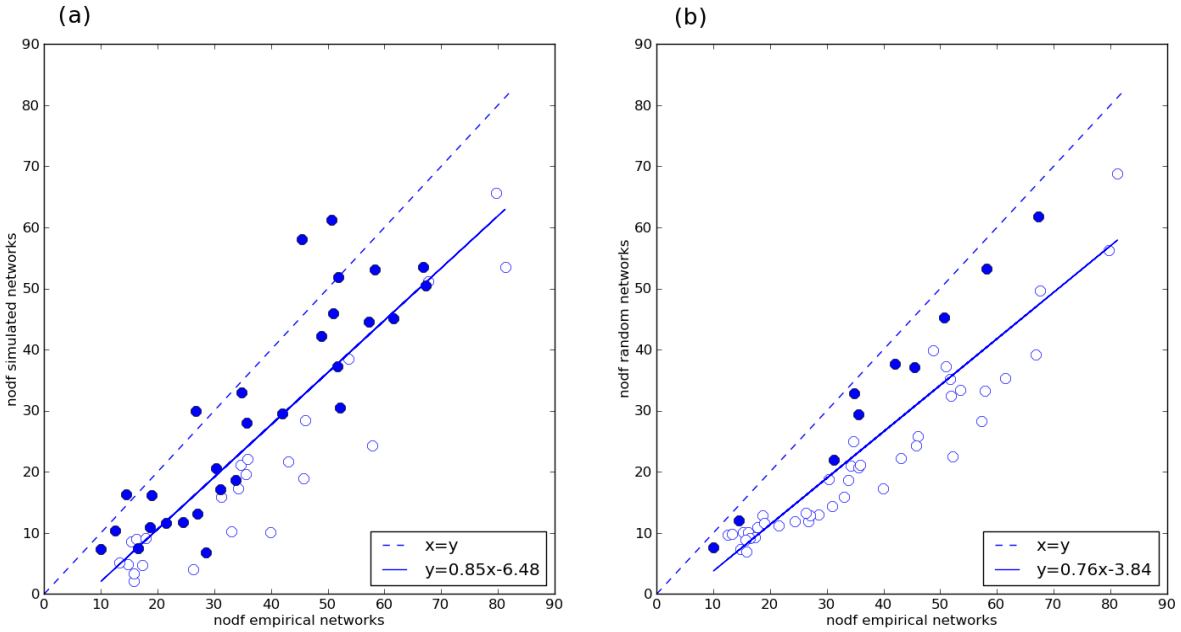

Figure 5.3: Average NODF values of networks simulated under our model (a) and of networks simulated under the null model (b) vs NODF values of empirical networks. Continuous lines are the best fits to the data, the broken lines represent x=y. Networks which discard the null hypotheses (our model or the null model for (a) or (b) respectively) are represented by empty circles. Otherwise, they are represented by solid circles.

5.3.2 Controlling False Discoveries

Some of the results considered significant (with a p-value ≤ 0.05), may in reality be mistakes. They are the false positive results. The proportion of false positive results can be controlled by fixing the false discovery rate (FDR) [Benjamini and Hochberg, 1995]. FDR is the expected percentage of false positives among all significant hypothesis. When the FDR is controlled to be less than a threshold of 0.05, our model is rejected for 41.51% of the networks and the null model is rejected for 79.24% of the networks.

Table 5.5: Absolute nestedness degrees

Networks	Empirical	Our model average±st.dev.	Null model average±st.dev.	p-value for our model	p-value for null model
ARR1	14.81	4.93±1.41	7.31±0.68	0.0	0.0
ARR2	15.28	8.64±1.59	10.1±1.18	0.0	0.0
ARR3	18.71	10.9±4.36	12.91±2.7	0.052	0.024
BAHE	33.73	18.6±12.04	18.69±2.63	0.142	0.0
BAIR	50.98	45.9±15.17	37.21±6.53	0.419	0.02
BEEH	67.66	51.24±6.65	49.65±4.5	0.004	0.0
CACG	46.0	28.5±9.09	25.86±4.35	0.017	0.0
CACI	43.0	21.74±9.3	22.19±3.34	0.003	0.0
CACO	30.34	20.55±8.63	18.83±4.3	0.134	0.007
CAFR	34.17	17.31±6.42	21.04±4.5	0.005	0.002
CLLO	15.81	2.05±0.89	6.93±0.43	0.0	0.0
CROM	51.74	37.21±18.89	35.24±3.88	0.31	0.0
DIHI	52.13	30.55±20.74	22.54±3.19	0.23	0.0
DISH	35.61	19.62±9.09	20.69±3.58	0.035	0.0
DUPO	35.64	28.04±6.39	29.41±3.72	0.135	0.058
EOL	16.19	9.01±1.4	10.16±1.08	0.0	0.0
EOLZ	34.65	21.2±3.01	24.99±1.38	0.0	0.0
ESKI	51.89	51.89±21.36	32.45±6.62	0.529	0.004
FROS	81.27	53.51±14.39	68.87±5.02	0.0	0.0
GEN1	34.82	33.01±9.59	32.81±7.22	0.415	0.386
GEN2	35.81	22.09±4.75	21.08±2.67	0.001	0.0
HAMM	50.68	61.29±14.94	45.33±3.32	0.697	0.054
HERR	21.47	11.59±7.37	11.27±1.19	0.142	0.0
HOCK	28.58	6.74±8.15	12.97±1.87	0.064	0.0
HRAT	79.75	65.68±8.94	56.23±4.29	0.003	0.0
INPK	17.82	9.13±1.68	11.01±1.04	0.0	0.0
KANT	67.34	50.45±14.84	61.85±4.97	0.059	0.137
KEVN	39.86	10.09±8.23	17.29±2.56	0.0	0.0
LOPE	58.22	53.17±9.47	53.23±5.95	0.325	0.193
MACK	10.03	7.39±2.77	7.66±1.76	0.173	0.094
MED1	18.92	16.16±14.72	11.6±2.42	0.301	0.004
MED2	24.42	11.78±12.16	11.85±2.06	0.138	0.0
MEMM	45.75	18.99±5.74	24.37±2.51	0.0	0.0
MOMA	31.24	15.85±6.57	21.9±5.54	0.015	0.059
MONT	32.96	10.29±3.5	15.91±1.0	0.0	0.0
MOTT	53.58	38.44±8.43	33.42±3.61	0.011	0.0
MULL	26.75	29.92±22.54	11.87±2.1	0.664	0.0
NCOR	57.21	44.5±14.86	28.32±2.96	0.296	0.0
NNOG	61.52	45.1±13.2	35.4±3.67	0.105	0.0
OLAU	31.0	17.13±8.81	14.45±1.95	0.091	0.0
PERC	27.03	13.08±12.22	12.89±1.97	0.123	0.0
PRAP	14.49	16.27±8.66	12.06±1.79	0.38	0.085
PRCA	17.28	4.77±2.18	9.31±0.8	0.0	0.0
PRCG	16.48	7.55±7.78	9.1±0.9	0.075	0.0
RABR	12.52	10.39±3.77	9.65±1.58	0.271	0.043
RMRZ	13.31	5.09±2.41	9.88±1.3	0.025	0.007
SAPF	66.8	53.47±18.23	39.24±7.56	0.279	0.0
SCHM	57.85	24.31±9.32	33.3±5.97	0.002	0.0
SMAL	42.03	29.57±13.35	37.71±3.4	0.244	0.092
SMRA	26.25	4.08±4.77	13.31±1.3	0.002	0.0
SNOW	48.81	42.24±4.51	39.94±2.66	0.078	0.0
WES	15.73	3.3±0.65	8.88±0.49	0.0	0.0
WYTH	45.41	58.09±16.44	37.16±6.37	0.751	0.091

Table 5.6: Relative nestedness degrees

Networks	Empirical	Our model average±st.dev.	Null model average±st.dev.	p-value for our model	p-value for null model
ARR1	1.027	-0.33±0.194	0.0±0.093	0.0	0.0
ARR2	0.513	-0.14±0.157	0.0±0.117	0.0	0.0
ARR3	0.449	-0.16±0.338	0.0±0.209	0.052	0.023
BAHE	0.804	-0.01±0.644	0.0±0.141	0.142	0.0
BAIR	0.37	0.23±0.408	0.0±0.176	0.419	0.02
BEEH	0.363	0.03±0.134	0.0±0.091	0.004	0.0
CACG	0.779	0.1±0.352	0.0±0.168	0.017	0.0
CACI	0.937	-0.02±0.419	0.0±0.151	0.003	0.0
CACO	0.611	0.09±0.458	0.0±0.228	0.134	0.007
CAFR	0.624	-0.18±0.305	0.0±0.214	0.005	0.002
CLLO	1.282	-0.7±0.128	0.0±0.063	0.0	0.0
CROM	0.468	0.06±0.536	0.0±0.11	0.31	0.0
DIHI	1.313	0.36±0.92	0.0±0.142	0.23	0.0
DISH	0.721	-0.05±0.439	0.0±0.173	0.035	0.0
DUPO	0.212	-0.05±0.217	0.0±0.126	0.135	0.058
EOL	0.593	-0.11±0.138	0.0±0.107	0.0	0.0
EOLZ	0.387	-0.15±0.12	0.0±0.055	0.0	0.0
ESKI	0.599	0.6±0.658	0.0±0.204	0.529	0.004
FROS	0.18	-0.22±0.209	0.0±0.073	0.0	0.0
GEN1	0.061	0.01±0.292	0.0±0.22	0.415	0.386
GEN2	0.699	0.05±0.225	0.0±0.126	0.001	0.0
HAMM	0.118	0.35±0.33	0.0±0.073	0.697	0.054
HERR	0.905	0.03±0.654	0.0±0.106	0.142	0.0
HOCK	1.203	-0.48±0.628	0.0±0.144	0.064	0.0
HRAT	0.418	0.17±0.159	0.0±0.076	0.003	0.0
INPK	0.619	-0.17±0.153	0.0±0.095	0.0	0.0
KANT	0.089	-0.18±0.24	0.0±0.08	0.057	0.136
KEVN	1.305	-0.42±0.476	0.0±0.148	0.0	0.0
LOPE	0.094	-0.0±0.178	0.0±0.112	0.324	0.191
MACK	0.31	-0.04±0.362	0.0±0.23	0.173	0.094
MED1	0.631	0.39±1.269	0.0±0.209	0.3	0.004
MED2	1.061	-0.01±1.027	0.0±0.174	0.138	0.0
MEMM	0.877	-0.22±0.235	0.0±0.103	0.0	0.0
MOMA	0.427	-0.28±0.3	0.0±0.253	0.015	0.059
MONT	1.072	-0.35±0.22	0.0±0.063	0.0	0.0
MOTT	0.603	0.15±0.252	0.0±0.108	0.011	0.0
MULL	1.254	1.52±1.899	0.0±0.177	0.664	0.0
NCOR	1.02	0.57±0.525	0.0±0.104	0.296	0.0
NNOG	0.738	0.27±0.373	0.0±0.104	0.105	0.0
OLAU	1.145	0.19±0.609	0.0±0.135	0.091	0.0
PERC	1.097	0.02±0.948	0.0±0.153	0.123	0.0
PRAP	0.202	0.35±0.718	0.0±0.148	0.38	0.084
PRCA	0.856	-0.49±0.234	0.0±0.086	0.0	0.0
PRCG	0.812	-0.17±0.855	0.0±0.099	0.075	0.0
RABR	0.298	0.08±0.39	0.0±0.163	0.27	0.042
RMRZ	0.347	-0.48±0.244	0.0±0.132	0.025	0.007
SAPF	0.702	0.36±0.465	0.0±0.193	0.279	0.0
SCHM	0.737	-0.27±0.28	0.0±0.179	0.002	0.0
SMAL	0.114	-0.22±0.354	0.0±0.09	0.244	0.092
SMRA	0.972	-0.69±0.359	0.0±0.098	0.002	0.0
SNOW	0.222	0.06±0.113	0.0±0.067	0.078	0.0
WES	0.772	-0.63±0.074	0.0±0.056	0.0	0.0
WYTH	0.222	0.56±0.442	0.0±0.171	0.751	0.091

5.4 Parameter Estimates

Estimates of the parameters we obtained by the maximum likelihood approach are given in table 5.7.

Values of the parameter λ (controlling the variation of branch lengths along the tree) that maximised the likelihood of the data are all ≤ 1. The optimized branches of the phylogenetic trees are all such that internal branches are shorter than external branches. Since those optimized branch lengths have been inferred from ecological data (the interaction networks), a high level of dissimilarity in the ecological data may explain long external branches: pairs of phylogenetically related species are more ecologically dissimilar than similar.

We did not find any pattern regarding the estimates of the parameter K (which scale the trees relative to each other). They are very diversified regarding each network.

Estimates of the gain-loss rate parameter (μ) range from 0.143 to 6.164 (with an average of 1.335).

For most of the networks (47 out of 53 networks), the existence of an interaction between two species appears to be less probable than the absence of an interaction ($\pi_0 > \pi_1$, (π_0, π_1) being the equilibrium frequency vector). Since we are using unique parameter μ for either the gain rate or the loss rate, it also appears that for those networks, as a species evolves, it loses an interaction at a higher rate (with a loss rate: $\mu\pi_0$) than it gains an interaction (with a gain loss: $\mu\pi_1$).

Taking into account the amount of evolutionary time since the most recent common ancestors of the species have existed (given by the total length of the phylogenetic trees), ecological changes (gain and loss events) have occurred on average 2.653 times per network.

Table 5.7: Maximum likelihood estimates of parameters in our stationary reversible model

Networks	λ	K	μ	equil. freq. π_0	gain rate	loss rate	amount of changes
ARR1	4	0.5	1.295	0.968	0.041	1.253	4.558
ARR2	4	0.25	3.075	0.93	0.216	2.859	10.827
ARR3	4	0.25	2.898	0.915	0.247	2.651	3.317
BAHE	2	3.5	0.228	0.855	0.033	0.195	2.398
BAIR	16	0.75	0.978	0.553	0.438	0.541	0.615
BEEH	32	0.25	2.954	0.555	1.315	1.639	4.347
CACG	16	0.25	3.441	0.817	0.629	2.812	4.231
CACI	8	0.25	2.35	0.893	0.252	2.098	1.789
CACO	4	0.25	2.294	0.85	0.345	1.95	3.071
CAFR	32	0.25	5.13	0.863	0.704	4.425	15.845
CLLO	8	2	1.261	0.984	0.02	1.241	4.561
CROM	8	3.5	0.367	0.662	0.124	0.243	1.361
DIHI	8	2.5	0.304	0.796	0.062	0.242	0.086
DISH	4	0.25	1.85	0.88	0.221	1.629	1.647
DUPO	2	2.5	0.403	0.745	0.103	0.3	0.463
EOL	4	0.25	2.197	0.913	0.192	2.005	3.089
EOLZ	4	0.25	1.87	0.844	0.291	1.579	6.585
ESKI	32	0.75	0.611	0.477	0.32	0.291	0.613
FROS	1	2.25	0.174	0.518	0.084	0.09	0.169
GEN1	32	0.25	6.164	0.678	1.986	4.178	16.734
GEN2	8	0.25	2.754	0.854	0.401	2.353	5.309
HAMM	1	0.75	0.185	0.43	0.106	0.08	0.118
HERR	2	2.5	0.275	0.915	0.023	0.251	0.103
HOCK	4	2.5	0.311	0.942	0.018	0.293	1.585
HRAT	16	0.5	1.254	0.393	0.761	0.493	1.441
INPK	8	0.25	3.333	0.932	0.225	3.108	3.039
KANT	8	2.5	0.469	0.457	0.255	0.214	1.241
KEVN	4	3.5	0.329	0.93	0.023	0.306	0.101
LOPE	1	0.75	0.506	0.511	0.247	0.259	3.544
MACK	16	0.25	5.938	0.936	0.381	5.556	5.96
MED1	2	2.5	0.143	0.888	0.016	0.127	0.041
MED2	4	2.5	0.268	0.918	0.022	0.246	0.055
MEMM	32	2.5	1.217	0.879	0.147	1.07	3.052
MOMA	2	2.5	0.838	0.849	0.127	0.711	1.126
MONT	8	2	0.745	0.931	0.051	0.694	0.309
MOTT	8	0.25	2.884	0.74	0.75	2.134	3.882
MULL	32	0.5	0.583	0.896	0.06	0.522	0.066
NCOR	8	2.5	0.27	0.703	0.08	0.19	1.139
NNOG	8	2.5	0.388	0.682	0.123	0.265	0.238
OLAU	2	3.5	0.212	0.907	0.02	0.192	3.339
PERC	4	1.25	0.342	0.903	0.033	0.309	1.186
PRAP	4	1	0.618	0.852	0.091	0.527	0.314
PRCA	4	3.5	0.547	0.962	0.021	0.526	0.174
PRCG	8	1	0.755	0.934	0.05	0.706	1.366
RABR	8	2.5	1.09	0.92	0.087	1.003	3.666
RMRZ	4	1.5	0.87	0.951	0.042	0.827	0.24
SAPF	4	2	0.272	0.491	0.139	0.134	0.981
SCHM	2	1	0.392	0.942	0.023	0.369	0.05
SMAL	2	1.5	0.314	0.736	0.083	0.231	0.173
SMRA	4	2.5	0.459	0.959	0.019	0.441	0.094
SNOW	2	0.25	1.186	0.656	0.408	0.778	5.338
WES	8	2.5	0.811	0.971	0.023	0.788	4.154
WYTH	8	1.75	0.403	0.449	0.222	0.181	0.875

5.5 Discussion

On the one hand, our model can evaluate an overall fit of the entire network while other models [Rezende et al., 2007a,b, Takemoto and Arita, 2010] only look at individual network properties in isolation to test the influence of phylogenetic history in shaping network structure. Comparison of the likelihood values given by the true phylogeny with likelihood values given by randomly generated phylogenies allowed us to do so. The likelihood comparison experiment also confirms the presence of a detectable phylogenetic signal in 88.68% of the networks.
On the other hand, our model also allows us to quantify the phylogenetic signal by looking at network properties in isolation:

1. Existence of already established models of node degree distribution allows us to use a model comparison approach for the node degree distribution study. We compare our model with those existing models. For about 34% of the networks, the shape of the node degree distribution is better explained by our model than by other models despite the fact that those other models are aimed to directly describe the node degree distribution and do not incorporate evolutionary processes. For the remaining networks, the information provided by the phylogeny is not enough to offset the advantages of using a direct model.

2. For nestedness study, only null models exist. This allows us to perform a hypothesis test: we tested whether our model can be rejected and whether the null model can be rejected and found that about 55% of the networks cannot reject our model while about only 19% of the networks cannot reject the null model. Even if the nestedness study does not permit us to quantify the proportion of networks which accept our model, cases where our model is rejected do not also indicate the presence of a better alternative.

From a global point of view, the results listed here suggest that our model performs substantially better than the null models without necessarily being the best performing model when considering specific network properties (degree distribution or nestedness) in isolation. Networks best fit by our model are most of the time small networks (composed of around 53 species). Although the effect was not significant (p-value=0.09), the degree of nestedness of a network may affect how well our model explains the degree distribution: it is possible that our model performs better for networks with high degree of nestedness. Since those networks have more structures in them, those structures are not easily lost along the evolu-

tionary time, compared to less structured networks.

We expect that networks which are well fit by our model should be the ones of which the properties reflect the past history of the species. For those networks, the inheritance of the state of each potential interaction from one generation to the next has been well described. Recall that in our model, the inheritance of the state of an interaction depends partially on the topology and the branch lengths of the phylogenetic trees (see 4.2.3). Consequently, dealing with the right phylogeny is crucial in our model. This has already been emphasized in the likelihood comparison study (see 5.1). Furthermore, dealing with phylogenetic trees with very long branch lengths results in random behaviour of the simulated networks. In fact, as time proceeds, phylogenetic distances also increase and phylogenetically related species tend less and less to share the same interaction, to a point of saturation. When saturation is reached, all phylogenetically valuable information is lost and the model will fit no better than if a random phylogeny were used. In the case where too short branch lengths are used, the model predictions will be too strong and will result in either very densely or very sparsely connected networks. The model also fits the data poorly for this case. The model performs correctly only with communities having phylogenetic trees with intermediate branch lengths. An illustration of this case study is given in figure 5.4.

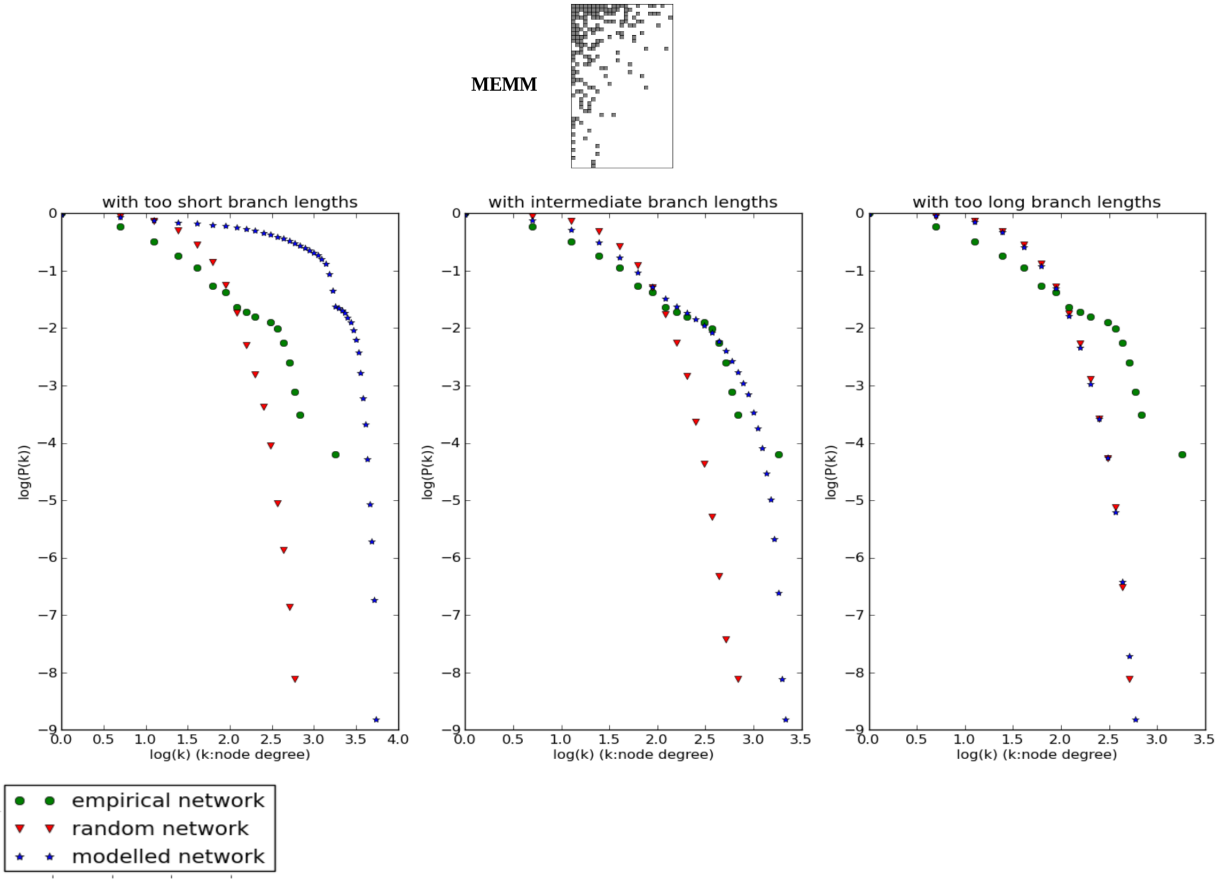

Figure 5.4: Log-log plot of the complementary cumulative distribution of the connectivities of an example of pollination network. In contrast to the inference using intermediate branch lengths, using too large branch lengths result in a random behaviour of the network while using too short branch lengths result in a poor fit to the data. The random networks (red triangle) were generated from the null model for the degree distribution study.

Using optimized branch lengths inferred from ecological data as a substitute for the true branch lengths leaves us with less information on the evolutionary history of the species, and might affect the detection of phylogenetic signal in some networks. Cases where branch lengths are not set to increase or decrease along the phylogenetic tree are for instance not considered in our simplified branch length model.

Our model is based on the evolutionary history: it assumes that descendant species also inherit ecological characteristics from parent species. This is realistic because morphological traits are inherited from parents to children. Since interactions between species depend on the complementarity of phenotypic traits of the species, when those interactions depend strongly on morphological traits, states of the interactions are more likely to be conserved. That is, when ecological change occurs rapidly compared to evolutionary change, ecological characteristics would not be conserved at the evolutionary time scale. In such cases we do not expect our model to fit well.

Nevertheless, our model is a simple evolutionary model that takes into account only the phylogenies of the species. This information may not be sufficient to model the entire evolutionary history. It may happen that for networks which are not well reproduced by our model, more complex evolutionary processes have occurred. Our model for instance neglects selective effects: some population of species may be affected by a selective pressure which can favour the evolution of beneficial interactions. This is not to say that structure of the networks badly fit by our model are inevitably not influenced by evolutionary history. A more elaborate model incorporating those evolutionary processes would be more adapted to detect phylogenetic signal in those networks.

Chapter 6

Conclusion

6.1 Summary

By assuming the theory of evolution stating that all organisms share a common ancestor, we have considered the phylogenetic trees relating the species involved in an ecological network as informative of their evolutionary history. Together with the state of the interaction network at present time, the phylogenetic history of the species has been incorporated in our model. In this manner we are able to obtain a reconstruction of the evolution through time of the ecological network involving ancestors of the current species. Our model is a very simple model. We assumed simplicity when building it: we made use of a stationary reversible model, and of biologically realistic parameters inferred from real data set.
While some ecologists have shown that the past history of the species involved in an interaction network plays a role in explaining some network patterns, our model allows us to confirm such finding: by analysing a large dataset composed of 53 pollinators and frugivory networks, the presence of a phylogenetic signal has been detected in about 89% of the networks. More than detecting the phylogenetic signal, we were able to quantify the extent to which evolutionary processes can explain these patterns.
Since the phylogenetic history of the species is the only factor influencing ecological network structure which is taken into consideration in our model, it makes sense that the structure of the networks simulated under our model are governed by only the phylogenetic history. Thus an analysis of those networks revealed that for a reasonably large subset of the networks (more than one third), taking the phylogenetic history into account gives a better explanation of node degree distribution than do models which are aimed to directly describe the node degree distribution. When instead the nestedness pattern of the networks is con-

sidered, we found that more than half of the networks cannot reject our model.
Even if our model can give an estimation of the importance of phylogenetic history in explaining network properties (node degree distribution and nestedness), its importance lies in the fact that it can give a description of the full network rather than just network properties in isolation. For this, to the best of our knowledge, no competing models to compare with our model can be found.

6.2 Recommendations and Perspectives

Since our model is a very simple model that incorporates simple evolutionary processes (only available information given by the phylogenetic trees), it can be considered as a null model for other more complicated models to test the significance of the importance of a particular evolutionary process in network structure.
Not having real branch lengths of the phylogenetic trees was an issue in our model. Even if we ended up using a model of branch lengths, all possible forms of a phylogenetic tree regarding the length of its branches could not be considered in our model whether the tree topology is fixed. Unfortunately, most available datasets do not include real branch lengths of the phylogenetic trees. We believe that if our model were used with dataset with branch lengths (as estimated from genetic or morphological data) rather than using our crude estimates informed by network structure, the results would be more accurate. Further investigations should be performed in this direction.
Although the case of mutualistic networks are emphasized in this thesis, the model we described here is adapted for any network such that the species can be divided into two distinct co-evolving groups and the network represented by a bipartite graph. We can for instance state the case of a herbivory network, which is an antagonistic network, but can be represented by a bipartite graph. Those cases are rare since most ecological bipartite networks are mutualistic networks. Furthermore, the method and the concepts we used when building the model can be easily adapted for general graphs.
Interaction matrices considered in our model are qualitative matrices in which each element corresponds to whether a presence or an absence of interaction. This implies considering an unweighted graph. Extension to quantitative matrices of which each element a_{ij} can for instance correspond to the number of visits between species i and species j also deserves attention. In that case, weighted bipartite graphs could instead be considered and a strength is assigned to each interaction.
The simple model we have built can be considered as a starting point for further research

incorporating phylogenetic analysis into ecological studies. We hope that this will lead to more elaborate and biologically realistic models in future.

Appendix A

Dataset

Table A.1: Details of each studied network

Networks code	Nature	Species number (animals/plants)	Locations
AARR1	pollination	181 (97/84)	Cordón del Cepo, Chile
ARR2	pollination	98 (60/38)	Cordón del Cepo, Chile
ARR3	pollination	50 (23/27)	Cordón del Cepo, Chile
BAHE	pollination	102 (91/11)	Central New Brunswick, Canada
BAIR	frugivory	28 (21/7)	Princeton, Mercer, New Jersey, USA
BEEH	frugivory	40 (9/31)	Mount Missim,Morobe Province, New Guinea
CACG	frugivory	38 (15/23)	Caguana, Puerto Rico
CACI	frugivory	53 (20/33)	Caguana, Puerto Rico
CACO	frugivory	36 (13/23)	Caguana, Puerto Rico
CAFR	frugivory	36 (15/21)	Fronton, Puerto Rico
CLLO	pollination	342 (246/96)	Pikes Peak, Colorado, USA
CROM	frugivory	77 (6/71)	Tropical rainforest, Queensland, Australia
DIHI	pollination	75 (58/17)	Hickling, Norfolk, UK
DISH	pollination	50 (34/16)	Shelfanger, Norfolk, UK
DUPO	pollination	46 (35/11)	Tenerife, Canary Islands
EOL	pollination	136 (112/24)	Latnjajaure, Abisko, Sweden
EOLZ	pollination	105 (74/31)	Zackenberg
ESKI	pollination	26 (12/14)	Mauritius Island
FROS	frugivory	24 (8/16)	Mtunzini, South Africa
GEN1	frugivory	24 (17/7)	Santa Genebra Reserve T1. SE Brazil

Continued on next page

GEN2	frugivory	61 (28/33)	Santa Genebra Reserve T2. SE Brazil
HAMM	frugivory	43 (16/27)	Negros Forest Reserve, Central Philippine Islands
HERR	pollination	190 (164/26)	Doñana Nat. Park, Spain
HOCK	pollination	100 (72/28)	Hazen Camp, Ellesmere Island, Canada
HRAT	frugivory	32 (16/16)	Hato Ratón, Sevilla, Spain
INPK	pollination	122 (80/42)	Snowy Mountains, Australia
KANT	frugivory	32 (27/5)	Campeche state, Mexico
KEVN	pollination	96 (76/20)	Hazen Camp, Ellesmere Island, Canada
LOPE	frugivory	25 (8/17)	Gabon, Africa
MACK	frugivory	60 (31/29)	Crater Mountain, Chimbu Province, Papua New Guinea
MED1	pollination	63 (43/20)	Laguna Diamante, Mendoza, Argentina
MED2	pollination	89 (68/21)	Rio Blanco, Mendoza, Argentina
MEMM	pollination	67 (42/25)	Bristol, England
MOMA	pollination	28 (17/11)	Melville Island, Canada
MONT	frugivory	206 (39/167)	Monteverde, Costa Rica
MOTT	pollination	55 (42/13)	North Carolina, USA
MULL	pollination	143 (39/104)	Galapagos
NCOR	frugivory	58 (33/25)	Nava Correhuelas.S. Cazorla, SE Spain
NNOG	frugivory	46 (28/18)	Nava Noguera, Sierra de Cazorla, SE Spain Flores
OLAU	pollination	83 (54/29)	Garajonay, Gomera, Spain
PERC	pollination	87 (31/56)	Jamaica
PRAP	pollination	71 (53/18)	Arthur's Pass, New Zealand
PRCA	pollination	172 (131/41)	Cass, New Zealand
PRCG	pollination	160 (111/49)	Craigieburn, New Zealand
RABR	pollination	72 (46/26)	Guarico State, Venezuela
RMRZ	pollination	91 (46/45)	Canaima National Park, Venezuela
SAPF	frugivory	23 (8/15)	Yakushima Island, Japan
SCHM	pollination	39 (32/7)	Brownfield, Illinois, USA
SMAL	pollination	45 (32/13)	Ottawa, Canada
SMRA	pollination	147 (121/26)	Chiloe, Chile
SNOW	frugivory	62 (14/48)	Tropical rainforest, Trinidad
WES	frugivory	286 (80/206)	Intervales and Saibadela, São Paulo, Brazil
WYTH	frugivory	25 (14/11)	Great Britain

References

G. Abramson, C. A. Trejo Soto, and L. Oña. The role of asymmetric interactions on the effect of habitat destruction in mutualistic networks. *PLoS ONE*, 6(6):e21028, 06 2011. doi: 10.1371/journal.pone.0021028.

M. Almeida Neto, P. Guimarães, P. R. Guimarães Jr, R. D. Loyola, and W. Ulrich. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, 117:1227–1239, march 2008. doi: 10.1111/j.2008.0030-1299.16644.x.

Jorge Aragón, David Eberly, and Shelly Eberly. Existence and uniqueness of the maximum likelihood estimator for the two-parameter negative binomial distribution. *Statistics & Probability Letters*, 15(5):375–379, December 1992.

Wirt Atmar and Bruce D. Patterson. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, 96(3):373–382, December 1993. doi: 10.1007/BF00317508.

Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen. The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9383–9387, August 2003. doi: 10.1073/pnas.1633576100.

J. Bascompte, P. Jordano, and J. M. Olesen. Asymmetric Coevolutionary Networks Facilitate Biodiversity Maintenance. *Science*, 312(5772):431–433, April 2006. doi: 10.1126/science.1123412.

Jordi Bascompte and Pedro Jordano. Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 38:567–593, 2007. doi: 10.1146/annurev.ecolsys.38.091206.095878.

U. Bastolla, M. A. Fortuna, A. Pascual-Garcia, A. Ferrera, B. Luque, and J. Bascompte. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458(7241):1018–1020, April 2009. ISSN 0028-0836. doi: 10.1038/nature07950.

Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995. ISSN 00359246. doi: 10.2307/2346101.

Simon P. Blomberg, Theodore Garland, and Anthony R. Ives. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution; international journal of organic evolution*, 57(4):717–745, April 2003. ISSN 0014-3820.

P. Borgnat, E. Fleury, J. L. Guillaume, C. Magnien, C. Robardet, and A. Scherrer. *Evolving Networks*, pages 198–204. NATO Science for Peace and Security Series D: Information and Communication Security. IOS Press, June 2008.

D. H. Boucher, S. James, and K. H. Keeler. The Ecology of Mutualism. *Annual Review of Ecology and Systematics*, 13:315–347, 1982.

J. L. Bronstein. Our current understand of mutualism. *Quarterly Review of Biology*, 69(1): 31–51, 1994.

Kenneth P. Burnham and David R. Anderson. Multimodel Inference. *Sociological Methods & Research*, 33(2):261–304, November 2004. ISSN 0049-1241. doi: 10.1177/0049124104268644.

Aaron Clauset, Cosma R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Reviews*, June 2007.

L. Da Fontoura Costa, O. N. Oliveira, Jr., G. Travieso, F. A. Rodrigues, R. V. B. Paulino, L. Antiqueira, M. P. Matheus, and L. E. C. da Rocha. Analyzing and Modeling Real-World Phenomena with Complex Networks: A Survey of Applications. *Advances in Physics*, 60: 329–412, 2011. doi: 10.1080/00018732.2011.572452.

Charles Darwin. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. *John Murray, London*, (1):1–556, October 1859.

L. V. Dicks, S. A. Corbet, and R. F. Pywell. Compartmentalization in plant-insect flower visitor webs. *Journal of Animal Ecology*, 71(1):32–43, 2002.

Y. L. Dupont, D. M. Hansen, and J. M. Olesen. Structure of a plant-flower-visitor network in the high-altitude sub-alpine desert of Tenerife, Canary Islands. *Ecography*, 26(3):301–310, 2003. doi: 10.1034/j.1600-0587.2003.03443.x.

Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, chapter 7. Cambridge University Press, July 1999a. ISBN 0521629713.

Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, chapter 1. Cambridge University Press, July 1999b. ISBN 0521629713.

J. Felsenstein. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of molecular evolution*, 17(6):368–376, 1981. ISSN 0022-2844.

Miguel A. Fortuna and Jordi Bascompte. Habitat loss and the structure of plantanimal mutualistic networks. *Ecology Letters*, 9(3):278–283, March 2006. ISSN 1461-023X. doi: 10.1111/j.1461-0248.2005.00868.x.

Nicholas J. Gotelli. Research frontiers in null model analysis. *Global Ecology & Biogeography*, 10(4):337–343, 2001. doi: 10.1046/j.1466-822X.2001.00249.x.

G. Grindrod and D. J. Higham. Models for evolving networks: with applications in telecommunication and online activities. *IMA Journal of Management Mathematics*, 2011. doi: 110.1093/imaman/dpr001.

P. Grindrod and D. J. Higham. Evolving graphs: dynamical models, inverse problems and propagation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 466(2115):753–770, 2010. URL `http://dx.doi.org/10.1098/rspa.2009.0456`.

P.R. Guimarães Jr, G. Machado, M.A.M. de Aguiar, P. Jordano, J. Bascompte, A. Pinheiro, and S.F. Dos Reis. Build-up mechanisms determining the topology of mutualistic networks. *Journal of Theoretical Biology*, Nov 2007. ISSN 0022-5193.

Jr Guimarães and Paulo Guimarães. Improving the analyses of nestedness for large sets of matrices. *Environmental Modelling & Software*, 21(10):1512–1513, October 2006. doi: 10.1016/j.envsoft.2006.04.002. URL `http://dx.doi.org/10.1016/j.envsoft.2006.04.002`.

P. R Guimarães, V. Rico-Gray, S. Furtado Dos Reis, and J. N. Thompson. Asymmetries in specialization in ant–plant mutualistic networks. *Proceedings of the Royal Society B Biological Sciences*, 273(1597):2041–2047, 2006.

M. Hasegawa, H. Kishino, and T. Yano. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution*, 22(2):160–174, 1985. ISSN 0022-2844.

L. N. Joppa, J. M. Montoya, R. Solé, J. Sanderson, and S. L. D. A. Pimm. On nestedness in ecological networks. *Evolutionary Ecology Research*, 12(1):35–46, 2010.

P. Jordano, J. Bascompte, and J. M. Olesen. Invariant properties in coevolutionary networks of plant animal interactions. *Ecology Letters*, 6:69–81, 2003.

P. Jordano, J. Bascompte, and J. M. Olesen. The ecological consequences of complex topology and nested structure in pollination webs. In N. M. Waser and J. Ollerton, editors, *Plant-Pollinator Interactions: From Specialization to Generalization*, pages 173–199. University Of Chicago Press, 2006.

J. Jost and M. P. Joy. Evolving networks with distance preferences. *Physical Review E*, 66, 2002. doi: 10.1103/PhysRevE.66.036126.

T. H. Jukes and C. R. Cantor. *Evolution of Protein Molecules*. Academy Press, 1969.

M. Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, (2):111–120, December 1980. ISSN 0022-2844.

S. L. Kosakovsky Pond, S. D. W. Frost, and S. V. Muse. HyPhy: hypothesis testing using phylogenies. 21(5):676–679+, 2005.

P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of Growing Random Networks. *Physical Review Letters*, 85:4629–4632, 2000.

C. L. Landry. Mighty Mutualisms: The Nature of Plant-pollinator Interactions. *Nature Education Knowledge*, 1(8):48, 2010.

D. J. Levey, W. R. Silva, and M. Galetti, editors. *Seed Dispersal and Frugivory: Ecology, evolution, and conservation*. CABI, 1st edition, April 2002. ISBN 085199525X.

T. M. Lewinsohn, Inacio, P. Jordano, J. Bascompte, and J. M. Olesen. Structure in plant-animal interaction assemblages. *Oikos*, 113(1):174–184, April 2006. doi: 10.1111/j.0030-1299.2006.14583.x.

Chunguang Li and Guanrong Chen. A comprehensive weighted evolving network model. *Physica A: Statistical Mechanics and its Applications*, 343:288–294, November 2004. doi: 10.1016/j.physa.2004.06.160.

Chunguang Li and Philip K. Maini. An evolving network model with community structure. *Journal of Physics A-mathematical and General*, 38:9741–9749, 2005. doi: 10.1088/0305-4470/38/45/002.

Xiang Li and Guanrong Chen. A local-world evolving network model. *Phisycal A: Statistical Mechanics and its Applications*, 328:274–286, 2003. doi: 10.1016/S0378-4371(03)00604-6.

Pietro Liò and Nick Goldman. Models of molecular evolution and phylogeny. *Genome Res*, 8:1233–1244, 1998.

D. Medan, R. P. Perazzo, M. Devoto, E. Burgos, M. G. Zimmermann, H. Ceva, and A. M. Delbue. Analysis and assembling of network structure in mutualistic systems. *Journal of Theoretical Biology*, 246:510–521, Jan 2007. doi: 10.1016/j.jtbi.2006.12.033.

Jane Memmott. The structure of a plant-pollinator food web. *Ecology Letters*, 2(5):276–280, January 1999.

Toshinori Okuyama. Do mutualistic networks follow power distributions? *Ecological Complexity*, 5:59–65, March 2008. doi: doi:10.1016/j.ecocom.2007.06.006.

J.M. Olesen, J. Bascompte, Y. L. Dupont, and P. Jordano. The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50):19891–19896, December 2007. doi: 10.1073/pnas.0706375104.

J. McCollin D. Fautin D. Ollerton and G. Allen. Finding nemo: nestedness engendered by mutualistic organization in anemonefish and their hosts. *Proceedings of the Royal Society Series B*, 274(1609), 2007. doi: 10.1098/rspb.2006.3758.

A. Onielsen and J. Bascompte. Ecological networks, nestedness and sampling effort. *Journal of Ecology*, 95, 2007. doi: 110.1111/j.1365-2745.2007.01271.x.

S. L. Pimm. Properties of food webs. *Ecology*, 61(2):219–225, 1980.

E. L. Rezende, P. Jordano, and J. Bascompte. Effects of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks. *Oikos*, 116(11):1919–1929, 2007a. ISSN 1600-0706. doi: 10.1111/j.0030-1299.2007.16029.x.

E. L. Rezende, J. E. Lavabre, P. R. Guimarães, P. Jordano, and J. Bascompte. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature*, 448(7156):925–928, August 2007b. ISSN 1476-4687. doi: 10.1038/nature05956.

Serguei Saavedra, Felix Reed-Tsochas, and Brian Uzzi. A simple model of bipartite cooperation for ecological and organizational networks. *Nature*, 457(7228):463–466, January 2009. ISSN 0028-0836. doi: 10.1038/nature07532.

H. L. Seal. The maximim likelihood fitting of the discrete pareto law. *Journal of the Institute of Actuaries*, 78, december 1952.

K. Takemoto and C. Oosawa. Modeling for evolving biological networks with scale-free connectivity, hierarchical modularity, and disassortativity. *Math Biosci*, 208, 2006. doi: 10.1016/j.mbs.2006.11.002.

Kazuhiro Takemoto and Masanori Arita. Nested structure acquired through simple evolutionary process. *Journal of Theoretical Biology*, 264(3):782–786, 2010.

S. Tavaré. *Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences*, volume 17, pages 57–86. Amer Mathematical Society, 1986. ISBN 0821811673.

W. Ulrich, M. Almeida-Neto, and N. J. Gotelli. A consumer's guide to nestedness analysis. *Oikos*, 118(1):3–17, 2009. ISSN 0030-1299. doi: 10.1111/j.1600-0706.2008.17053.x.

A. Vázquez, A. Flammini, A. Maritan, and A. Vespignani. Modeling of protein interaction networks. *Complexus*, 1:38–44, 2003. doi: 10.1159/000067642. URL `http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:cond-mat/0108043`.

D. P. Vázquez, N. Bluthgen, L. Cagnolo, and N. P. Chacoff. Uniting pattern and process in plant-animal mutualistic networks: a review. *Ann Bot*, 103(9):1445–1457, June 2009. doi: 10.1093/aob/mcp057.

Diego P. Vázquez. Degree distribution in plant-animal mutualistic networks: forbidden links or random interactions? *Oikos*, 108:421+, 2005. ISSN 0030-1299.

Diego P. Vázquez and Marcelo A. Aizen. Asymmetric specialization: a pervasive feature of plant-pollinator interactions. *Ecology*, 85:1251–1257, 2004.

Diego P. Vázquez and Marcelo A. Aizen. Null model analyses of specialization in plant-pollinator interactions. *Ecology*, 84(9):2493–2501, 2003.

N. M. Waser, L. Chittka, M. V. Price, N. M. Williams, and J. Ollerton. Generalization in Pollination Systems, and Why it Matters. *Ecology*, 77(4):1043–1060, 1996. ISSN 00129658. doi: 10.2307/2265575. URL `http://dx.doi.org/10.2307/2265575`.

Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998. ISSN 0028-0836. doi: 10.1038/30918.

R. J. Williams and N. D. Martinez. Simple rules yield complex food webs. *Nature*, 404 (6774):180–183, March 2000. doi: 10.1038/35004572.